



TUGAS AKHIR - SS141501

**PENGELOMPOKAN KABUPATEN/ KOTA DI JAWA
TIMUR BERDASARKAN PEMBANGUNAN KUALITAS
SUMBER DAYA MANUSIA DAN PEMBANGUNAN
EKONOMI DENGAN METODE *ENSEMBLE* ROCK**

**LILIK INDRIYATI ICHSAN
NRP 062114 4000 0040**

**Dosen Pembimbing
Erma Oktania Permatasari, S.Si., M.Si
Dr. Vita Ratnasari, S.Si., M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

**PENGELOMPOKAN KABUPATEN/ KOTA DI JAWA
TIMUR BERDASARKAN PEMBANGUNAN KUALITAS
SUMBER DAYA MANUSIA DAN PEMBANGUNAN
EKONOMI DENGAN METODE *ENSEMBLE ROCK***

**LILIK INDRIYATI ICHSAN
NRP 062114 4000 0040**

**Dosen Pembimbing
Erma Oktania Permatasari, S.Si., M.Si
Dr. Vita Ratnasari, S.Si., M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS141501

***THE CLASSIFICATION OF REGION/ DISTRICT
IN EAST JAVA BASED ON THE DEVELOPMENT
OF HUMAN RESOURCE AND ECONOMIC
DEVELOPMENT USING ENSEMBLE ROCK
METHOD***

**LILIK INDRIYATI ICHSAN
SN 062114 4000 0040**

Supervisor

**Erma Oktania Permatasari, S.Si., M.Si
Dr. Vita Ratnasari, S.Si., M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

PENGELOMPOKAN KABUPATEN/ KOTA DI JAWA TIMUR BERDASARKAN PEMBANGUNAN SUMBER DAYA MANUSIA DAN PEMBANGUNAN EKONOMI DENGAN METODE *ENSEMBLE ROCK*

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

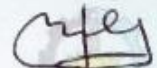

Oleh :

Lilik Indriyati Ichsan
NRP. 06211440000040

Disetujui oleh Pembimbing :

Erma Oktania Permatasari, S.Si., M.Si
NIP. 19881007 201404 2 002

Dr. Vita Ratnasari, S.Si., M.Si
NIP. 19700910 199702 2 001

()
()

Mengetahui
Kepala Departemen



Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2018

(Halaman ini sengaja dikosongkan)

**PENGELOMPOKAN KABUPATEN/ KOTA DI JAWA
TIMUR BERDASARKAN PEMBANGUNAN SUMBER
DAYA MANUSIA DAN PEMBANGUNAN EKONOMI
DENGAN METODE *ENSEMBLE ROCK***

Nama : Lilik Indriyati Ichsan
NRP : 062114 4000 0040
Departemen : Statistika
Dosen Pembimbing : Erma Oktania P., S.Si., M.Si
Vita Ratnasari, S.Si., M.Si

Abstrak

Pada tahun 2020-2030 mendatang Indonesia akan menghadapi peristiwa bonus demografi. Adanya bonus demografi diharapkan mampu menjadi angin segar bagi pemerintah dalam upaya mengentaskan masalah kemiskinan. Dalam rangka mencapai tujuan tersebut dan menjadikan peristiwa bonus demografi sebagai the window of opportunity diperlukan kualitas SDM (Sumber Daya Manusia) yang memadai dan kebijakan pemerintah yang sejalan dengan potensi yang ada. Provinsi Jawa Timur merupakan provinsi yang memiliki jumlah penduduk terbesar nomor dua di Indonesia dan jumlah kemiskinan serta kesenjangan sosial yang masih tergolong tinggi. Variabel yang digunakan dalam penelitian ini merupakan variabel yang menjadi indikator penilaian pembangunan kualitas SDM dan pembangunan ekonomi yang terdiri dari 21 variabel campuran numerik dan kategori. Sehingga metode pengelompokan yang digunakan untuk menyelesaikan persoalan tersebut adalah metode ensemble ROCK. Hasil dari pengelompokan dengan metode ensemble ROCK menunjukkan bahwa kelompok optimum yang terbentuk adalah sebanyak dua kelompok dengan threshold (θ) sebesar 0.05 dan nilai rasio sebesar 2.98×10^{-16} .

Kata Kunci : *Ekonomi, Ensemble ROCK, Kategori, Kualitas SDM, Numerik*

(Halaman ini sengaja dikosongkan)

***THE CLASSIFICATION OF REGION/ DISTRICT IN EAST
JAVA BASED ON THE DEVELOPMENT OF HUMAN
RESOURCE AND ECONOMIC DEVELOPMENT USING
ENSEMBLE ROCK METHOD***

Name : Lilik Indriyati Ichsan
Student Number : 062114 4000 0040
Department : Statistics
Supervisor : Erma Oktania P., S.Si., M.Si
Vita Ratnasari, S.Si., M.Si

Abstract

In the next 2020-2030, Indonesia will face the demographic bonus event. The existence of demographic bonus is expected to be a fresh breeze for the government in an effort to alleviate the problem of poverty. In order to achieve these objectives and make demographic bonus events as the window of opportunity requires adequate quality of human resources (HR) and government policies that are in line with existing potential. East Java Province was selected in this study because this province has the second largest population in Indonesia and the number of poverty and social gaps is still relatively high. The variables used in this study are the variables that become indicators of assessment of the development of human resource quality and economic development consisting of 21 variables of numerical mix and category. So the grouping method used to solve the problem is the ROCK ensemble method. The result of clustering using ROCK ensemble method showed that the optimum group was formed as two groups with the threshold of 0.05 and the value of ratio was 2.98×10^{-16} .

Keywords : *Category, Economics, Ensemble ROCK, Numeric, Quality of Human Resources*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji dan syukur dipanjatkan kepada Allah SWT, Tuhan semesta alam. Atas segala rahmat dan nikmat-Nya yang diberikan kepada penulis, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Pengelompokan Kabupaten/ Kota Berdasarkan Pembangunan Sumber Daya Manusia dan Pembangunan Ekonomi dengan Metode *Ensemble ROCK*”** dengan baik dan tepat pada waktunya.

Penyusunan Tugas Akhir ini tidak lepas dari bantuan serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada :

1. Orang tua dan keluarga yang selalu memberikan dukungan baik secara moral maupun materi demi terselesaikannya Tugas Akhir ini.
2. Bapak Dr. Suhartono selaku Kepala Departemen Statistika ITS dan Bapak Dr. Sutikno, M.Si selaku Ketua Program Studi Sarjana Departemen Statistika ITS yang telah memberikan fasilitas untuk kelancaran pengerjaan Tugas Akhir.
3. Ibu Erma Oktania Permatasari, S.Si., M.Si dan Ibu Dr. Vita Ratnasari, S.Si., M.Si selaku dosen pembimbing yang dengan sabar telah memberikan ilmu, saran, dan nasihat selama penyusunan Tugas Akhir.
4. Bapak Prof. Dr. I Nyoman Budiantara, M.Si dan Dr. Agnes Tuti Rumiati, MSc selaku dosen penguji yang telah memberikan kritik dan saran demi kesempurnaan Tugas Akhir ini.
5. Bapak R. Moh Atok, S.Si., M.Si selaku dosen wali yang senantiasa memberikan nasihat dan motivasi selama masa perkuliahan.
6. Teman-teman angkatan 2014 yang sama-sama berjuang dalam menyelesaikan Tugas Akhir dan saling memberikan semangat satu dengan yang lain.
7. Sahabat penulis, Winda Dwi Asmianti dan Khurnia Risqi Nur Latifa yang selalu mengingatkan dan memberikan semangat

kepada penulis agar dapat menyelesaikan Tugas Akhir ini tepat waktu.

Semoga kebaikan-kebaikan yang telah diberikan kepada penulis dibalas oleh Allah SWT dengan kebaikan yang lebih besar lagi. Akhir kata penulis berharap dengan adanya Tugas Akhir ini dapat bermanfaat bagi pembaca dan penelitian selanjutnya.

Surabaya, Juli 2018

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	6
1.3 Tujuan Penelitian	7
1.4 Manfaat Penelitian	7
1.5 Batasan Masalah	7
BAB II TINJAUAN PUSTAKA	9
2.1 Data Kategori dan Data Numerik	9
2.2 Analisis Multivariat	9
2.3 Analisis Klaster	10
2.4 Analisis Klaster Data Numerik	12
2.4.1 Klaster Metode Hierarki	13
2.4.2 Klaster Metode Non Hierarki	14
2.5 Analisis Klaster Data Kategori	15
2.6 Analisis Klaster Data Campuran	18
2.7 Kinerja Hasil Pengelompokan	20
2.7.1 Kinerja Hasil Pengelompokan Data Numerik	20
2.7.2 Kinerja Hasil Pengelompokan Data Kategori	22
2.8 Pembangunan Nasional	23
2.8.1 Pembangunan SDM	24

2.8.2	Pembangunan Ekonomi.....	25
2.9	Gambaran Umum Provinsi Jawa Timur	25
2.10	Penelitian Sebelumnya.....	26
2.11	Kerangka Konsep Penelitian.....	28
BAB III	METODOLOGI PENELITIAN.....	29
3.1	Sumber Data	29
3.2	Variabel Penelitian.....	29
3.3	Langkah Analisis	33
BAB IV	ANALISIS DAN PEMBAHASAN.....	37
4.1	Karakteristik Kabupaten/ Kota di Jawa Timur Berdasarkan Pembangunan Kualitas SDM dan Pembangunan Ekonomi	37
4.1.1	Karakteristik Berdasarkan Skala Data Numerik....	37
4.1.2	Karakteristik Berdasarkan Skala Data Kategori....	40
4.2	Pengelompokan Kabupaten/ Kota di Jawa Timur Berdasarkan Pembangunan Kualitas SDM dan Pembangunan Ekonomi	43
4.2.1	Pengelompokan Data Numerik	44
4.2.2	Pengelompokan Data Kategori.....	45
4.2.3	Pengelompokan Data Campuran	46
4.2.4	Karakteristik Setiap Kelompok	50
BAB V	KESIMPULAN DAN SARAN.....	57
5.1	Kesimpulan	57
5.2	Saran	58
	DAFTAR PUSTAKA	61
	LAMPIRAN	65

DAFTAR GAMBAR

Gambar 2.1 Tahapan Umum Pengelompokan <i>Ensemble</i>	19
Gambar 2.2 Tahapan Umum Algoritma CEBMDC	20
Gambar 2.3 Struktur PDRB Jawa Timur Tahun 2016.....	26
Gambar 2.4 Kerangka Konsep Penelitian.....	28
Gambar 3.1 Diagram Alir Penelitian	35
Gambar 4.1 Diagram Lingkaran Indeks Gini Ratio.....	41
Gambar 4.2 Diagram Lingkaran Laju Pertumbuhan Ekonomi ..	42
Gambar 4.3 Peta Daerah Ekonomi	42
Gambar 4.4 Diagram Batang Sektor Penyumbang PDRB Terbesar	43
Gambar 4.5 Peta Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Hasil Pengelompokan <i>Ensemble</i> ROCK	55

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

Tabel 2.1	Struktur Data Multivariat	9
Tabel 2.2	Tabel Kontingensi Data Biner	11
Tabel 2.3	Ukuran Jarak Data Biner	11
Tabel 3.1	Variabel Penelitian Berskala Numerik	29
Tabel 3.2	Variabel Penelitian Berskala Kategori	30
Tabel 4.1	Karakteristik Data Numerik	37
Tabel 4.2	Hasil Kinerja Pengelompokan Data Numerik	44
Tabel 4.3	Hasil Kinerja Pengelompokan Data Kategori	45
Tabel 4.4	Rangkuman Hasil Pengelompokan Data Numerik dan Kategori	46
Tabel 4.5	Struktur Data Baru	48
Tabel 4.6	Hasil Kinerja Pengelompokan Data Campuran	49
Tabel 4.7	Daftar Anggota Hasil Pengelompokan Data Campuran	50
Tabel 4.8	Karakteristik Kelompok Satu Berdasarkan Data Numerik	50
Tabel 4.9	Karakteristik Kelompok Dua Berdasarkan Data Numerik	51
Tabel 4.10	Perbandingan Nilai <i>Mean</i> Setiap Kelompok	52
Tabel 4.11	Karakteristik Setiap Kelompok Berdasarkan Data Kategori	52

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1	Data Berskala Numerik	65
Lampiran 2	Data Berskala Kategori	68
Lampiran 3	<i>Syntax</i> untuk Menghitung Nilai <i>Icdrate</i>	70
Lampiran 4	<i>Syntax</i> Pengelompokan Data Kategori	72
Lampiran 5	<i>Syntax</i> Pengelompokan Data Campuran	79
Lampiran 6	Anggota Pengelompokan Data Numerik	86
Lampiran 7	Anggota Pengelompokan Data Kategori	88
Lampiran 8	Anggota Pengelompokan Data Campuran	90
Lampiran 9	Hasil Pengelompokan Data Kategori dengan <i>Threshold</i> Lainnya	92
Lampiran 10	Hasil Pengelompokan Data Campuran dengan <i>Threshold</i> Lainnya	93
Lampiran 11	Surat Pernyataan Pengambilan Data	94

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Negara anggota PBB (Perserikatan Bangsa Bangsa) telah sepakat untuk menciptakan dunia yang lebih sejahtera, adil, dan damai. Kesepakatan tersebut yang menjadi awal dari pembentukan tujuan pembangunan berkelanjutan atau yang lebih dikenal dengan MDGs (*Millenium Development Goals*). MDGs yang selesai pada tahun 2015 lalu, kemudian digantikan dengan SDGs (*Sustainable Development Goals*) yang telah dimulai sejak tahun 2016 hingga 2030 mendatang. SDGs memiliki tujuan yang hampir sama dengan MDGs, yakni penghapusan kemiskinan, penghapusan kelaparan, kesehatan dan kesejahteraan, pendidikan berkualitas, kesetaraan gender, air bersih dan sanitasi, energi bersih dan terjangkau, pertumbuhan ekonomi dan pekerjaan yang layak, infrastruktur tangguh, penurunan kesenjangan, kota inklusif berkelanjutan, konsumsi dan produksi berkelanjutan, perubahan iklim dan pengurangan resiko bencana, pelestarian dan pemanfaatan berkelanjutan ekosistem laut, pelestarian dan pemanfaatan ekosistem darat, perdamaian, serta kemitraan untuk semua tujuan pembangunan (UNDP, 2017).

Salah satu tujuan yang sampai saat ini masih belum tercapai adalah penghapusan kemiskinan. Kemiskinan merupakan salah satu masalah yang masih banyak terjadi di negara-negara berkembang, termasuk di Indonesia. Adanya permasalahan kemiskinan yang dikenal sebagai lingkaran setan, menyebabkan tingkat pendidikan dan kesehatan masyarakat menjadi rendah. Tingkat pendidikan dan kesehatan yang rendah menjadi penyebab utama rendahnya tingkat produktivitas dan kesejahteraan masyarakat. Berbentuk negara kepulauan menjadikan pemerataan kesejahteraan di Indonesia cukup sulit dicapai. Kesenjangan sosial banyak terjadi di daerah-daerah atau pulau-pulau yang jauh dari Ibu Kota. Fakta menunjukkan bahwa masyarakat yang tinggal jauh dari ibu kota mengalami kesulitan dalam berbagai hal, bahkan untuk

mengakses kebutuhan yang paling mendasar seperti kesehatan dan pendidikan. Sehingga tidak jarang daerah-daerah yang jauh dari Ibu Kota apalagi di luar pulau Jawa terisolasi dan tidak mempunyai akses terhadap sumber daya.

Indonesia merupakan negara yang terkenal dengan SDA (Sumber Daya Alam) yang melimpah. Mulai dari sektor perikanan, pertanian, perkebunan, hingga pertambangan ada di Indonesia. Namun, hal tersebut ternyata belum sepenuhnya dimanfaatkan dengan baik oleh penduduk Indonesia. Tidak jarang perusahaan-perusahaan yang mengelola kekayaan SDA di Indonesia merupakan perusahaan asing. Pemerintah tidak boleh tinggal diam melihat SDA yang dimiliki justru dikuasai oleh orang asing. Sementara itu, Indonesia juga tidak bisa terus menerus mengandalkan SDA yang dimiliki, karena tidak semua SDA bersifat kekal. Akan tiba masanya beberapa SDA di Indonesia habis terpakai atau bahkan rusak karena tidak dikelola dengan baik. Adanya SDM (Sumber Daya Manusia) yang mumpuni dan berkualitas bisa menjadi penyelamat suatu negara ketika mengalami krisis SDA. Secara teori, kualitas sumber SDM menjadi kunci utama dalam pertumbuhan ekonomi berkelanjutan dan berkeadilan.

Berdasarkan publikasi yang dikeluarkan oleh UNFPA (2010) menyatakan bahwa pada tahun 2020 – 2030 Indonesia akan mengalami fenomena bonus demografi. Hal ini menjadi momentum yang sangat penting bagi pemerintah untuk lebih giat lagi melakukan usaha-usaha dalam meningkatkan kualitas SDM yang dimiliki. Bonus demografi bisa menjadi *the window of opportunity* bagi Indonesia apabila didukung dengan kualitas SDM yang baik. Adanya SDM yang berkualitas dapat meningkatkan produktivitas masyarakat serta berpengaruh terhadap kesejahteraan dan pertumbuhan ekonomi. Namun, disisi lain bonus demografi juga bisa menjadi momentum pengangguran besar-besaran apabila jumlah angkatan kerja yang ada, tidak sebanding dengan jumlah lapangan kerja yang tersedia. Dalam pidatonya saat menghadiri penyerahan sertifikat Kompetensi Peserta Pemagangan tahun 2017 di Balai Besar Pengembangan Latihan Kerja (BBPLK) di Bekasi,

Presiden Jokowi mengatakan bahwa pada tahun 2019, pembangunan akan difokuskan pada pembangunan SDM (Ihsanuddin, 2017). Oleh karena itu, pada penelitian ini peneliti ingin melakukan pengelompokan terhadap kabupaten/kota di Provinsi Jawa Timur berdasarkan pembangunan kualitas SDM dan pembangunan ekonomi. Hasil dari pengelompokan tersebut bertujuan untuk mengetahui dampak dari program pembangunan yang telah dilakukan oleh pemerintah hingga saat ini pada masing-masing kelompok yang terbentuk. Berdasarkan informasi pada masing-masing kelompok tersebut diharapkan dapat membantu fokus pembangunan SDM pada tahun 2019 mendatang agar lebih tepat sasaran.

Provinsi Jawa Timur merupakan provinsi dengan jumlah penduduk terbanyak nomor dua di Indonesia. Hal ini tentu menjadi pekerjaan rumah yang cukup berat bagi pemerintah. Selain harus mengatasi masalah kemiskinan dan kesenjangan sosial yang terjadi, peningkatan kualitas SDM juga harus segera dilakukan agar Provinsi Jawa timur dapat menikmati adanya peristiwa bonus demografi.

Sesuai dengan pasal 1 ayat 22 pada Undang-Undang Nomor 23 Tahun 2017 tentang pemerintahan daerah, cakupan wilayah provinsi adalah daerah kabupaten/kota. Daerah kabupaten/kota merupakan daerah yang masing-masing memiliki pemerintahan daerah, sehingga bupati/walikota memiliki otonomi untuk mengatur daerahnya sendiri. Selain itu, pembagian daerah ekonomi dalam penyelenggaraan penyiaran untuk kabupaten/kota telah diatur dalam Peraturan Menteri Komunikasi dan Informatika Republik Indonesia Nomor 39 Tahun 2008. Dalam peraturan tersebut menyatakan bahwa daerah ekonomi maju merupakan daerah yang diklasifikasikan pada zona 1, zona 2, dan zona 3. Sedangkan zona 4 dan zona 5 merupakan daerah ekonomi kurang maju. Pembagian daerah ekonomi kemudian dimasukkan oleh peneliti dalam variabel input untuk membedakan setiap kabupaten/kota berdasarkan klasifikasi tersebut.

Dr. Soni Sumarsono, M. DM selaku Direktur Jenderal Otonomi Daerah Kementerian Dalam Negeri yang hadir sebagai

pembicara pada Workshop Demography Forum UGM menjelaskan bahwa salah satu upaya yang akan dilakukan oleh Kemendagri dalam upaya pemanfaatan bonus demografi adalah pemetaan keahlian penduduk usia produktif. Pemetaan dilakukan pada sektor yang kontributif pada pertumbuhan ekonomi agar nantinya tenaga kerja lebih tanggap dalam merespon kebutuhan pasar (Satria, 2017). Kebutuhan Indonesia untuk saat ini adalah memanfaatkan potensi terbaik untuk pembangunan perekonomian yang berkelanjutan dalam jangka panjang. Sinergi kebijakan antar pemangku kepentingan pada sektor terkait dan lintas sektor juga mutlak diperlukan untuk menyatukan sumber daya dan potensi yang ada bagi percepatan pembangunan SDM di Indonesia. Hal inilah yang menjadi landasan peneliti mengkategorikan PDRB berdasarkan sektor yang paling mendominasi pada tahun 2016. Selain itu, peneliti mengkategorikan variabel indeks *gini ratio* dan laju pertumbuhan ekonomi berdasarkan pencapaian pada target yang tercantum pada APBN 2016. Pengkategorian dilakukan agar dapat diketahui daerah-daerah yang sudah mencapai target pembangunan dan daerah-daerah yang belum mencapai target pembangunan. Penggunaan variabel kategori tersebut, mulai dari pembagian daerah ekonomi hingga laju pertumbuhan ekonomi diharapkan mampu menjadikan hasil pengelompokan yang terbentuk lebih mudah untuk dipahami. Pada penelitian ini peneliti juga menggunakan variabel numerik yang berupa indikator penyusun IPM. Indikator penyusun IPM dipilih karena IPM merupakan cerminan dari pembangunan kualitas SDM di suatu wilayah. Total variabel yang digunakan dalam penelitian ini adalah sebanyak 21 variabel, dengan rincian 17 variabel merupakan variabel data numerik dan 4 variabel merupakan variabel data kategori.

Penelitian mengenai pengelompokan menggunakan metode analisis kluster telah banyak dilakukan sebelumnya. Pravitasari (2008) melakukan pengelompokan kecamatan dengan menggunakan metode *Fuzzy C-Means* yang merupakan pengembangan dari metode kluster non hierarki. Sementara itu, Arianti (2009)

mengelompokkan kecamatan di Kabupaten Probolinggo berdasarkan Indeks Pembangunan Manusia dengan metode kluster hierarki. Dua tahun kemudian, Lailiya (2011) membandingkan metode kluster hierarki dan metode kluster non hierarki pada pengelompokan kabupaten/kota di Jawa Timur berdasarkan faktor-faktor yang mempengaruhi TPT. Hasil pengelompokan terbaik didapatkan oleh pengelompokan metode kluster non hierarki *Fuzzy C-Means* dengan kelompok optimum sebanyak 6. Namun, metode ini tidak cocok jika digunakan dalam penelitian ini karena pada data yang ada tidak memiliki informasi awal mengenai jumlah kelompok yang terbentuk.

Dewi (2012) melakukan penelitian mengenai metode kluster *ensemble* untuk pengelompokan desa perdesaan di Provinsi Riau. Tujuan dari penelitian tersebut adalah mengkaji dan membangun program untuk metode pengelompokan *ensemble* dalam menangani variabel berskala campuran numerik dan kategori. Hasilnya menunjukkan bahwa metode pengelompokan *ensemble* menghasilkan kelompok dengan kinerja yang lebih baik dari pada pengelompokan *full* kategori dan *full* numerik. Kemudian Alvionita (2017) melakukan penelitian yang membandingkan metode *ensemble* ROCK dan SWFM untuk pengelompokan data campuran pada kasus aksesori jeruk. Berdasarkan penelitian tersebut metode ROCK memberikan kinerja pengelompokan lebih baik karena menghasilkan rasio antara simpangan baku di dalam kelompok dan simpangan baku antar kelompok yang lebih kecil dari pada metode SWFM. Di tahun yang sama, Permata (2017) melakukan pengelompokan PTN (Perguruan Tinggi Negeri) di Indonesia menggunakan metode *ensemble* ROCK (*RObust Clustering using linKs*). Pada pengelompokan menggunakan metode ROCK diperoleh 2 kelompok, yaitu PTN dengan kualitas rendah dan PTN dengan kualitas tinggi.

Adanya variabel campuran, yaitu kategori dan numerik dalam penelitian ini membuat pengelompokan tidak bisa dilakukan dengan metode analisis kluster biasa. Diperlukan metode khusus untuk menangani pengelompokan data campuran. Oleh karena itu,

pada penelitian ini peneliti memutuskan untuk menggunakan metode *ensemble* ROCK dalam penyelesaian persoalan pengelompokan data campuran. Asumsi yang digunakan peneliti adalah berdasarkan penelitian sebelumnya yang menyatakan bahwa metode *ensemble* ROCK menghasilkan kinerja pengelompokan yang lebih baik dari pada metode SWFM. Proses analisis dimulai dengan mengelompokkan data murni numerik menggunakan metode klaster hierarki *agglomerative*. Metode *agglomerative* yang digunakan adalah *Single Linkage Method*, *Complete Linkage Method*, *Average Linkage*, dan *Ward's Method*. Berdasarkan keempat metode hierarki tersebut kemudian dipilih hasil pengelompokan terbaik dengan membandingkan nilai *icdrate* yang diperoleh. Sedangkan data murni kategori akan dikelompokkan dengan menggunakan metode ROCK. Kemudian kedua hasil pengelompokan tersebut dikelompokkan kembali menggunakan metode *ensemble* ROCK untuk mendapatkan *final cluster*.

1.2 Rumusan Masalah

Pembangunan adalah proses perubahan yang mencakup seluruh sistem sosial, seperti politik, ekonomi, infrastruktur, pertahanan, pendidikan dan teknologi, kelembagaan, serta budaya. Dalam pidatonya, Presiden Jokowi mengatakan bahwa pada tahun 2019 konsentrasi pemerintah akan difokuskan pada pembangunan SDM untuk pemanfaatan bonus demografi. Di sisi lain, kualitas SDM yang dimiliki suatu daerah akan mempengaruhi pertumbuhan ekonomi daerah tersebut. Berdasarkan permasalahan tersebut, maka perlu dilakukan pengukuran keberhasilan pembangunan nasional untuk menentukan kebijakan yang tepat untuk mencapai target pembangunan. Permasalahan utama yang akan dibahas pada penelitian ini adalah pengelompokan kabupaten/ kota di Provinsi Jawa Timur menggunakan metode pengelompokan *ensemble* ROCK. Pengelompokan tersebut dilakukan berdasarkan pembangunan nasional dengan skala data campuran numerik dan kategori.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka tujuan yang ini dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengetahui gambaran umum mengenai kondisi pembangunan kualitas SDM dan pembangunan ekonomi di Provinsi Jawa Timur pada tahun 2016.
2. Mendapatkan kelompok untuk kabupaten/kota di Provinsi Jawa Timur berdasarkan pembangunan kualitas SDM dan pembangunan ekonomi dengan metode *ensemble* ROCK.

1.4 Manfaat Penelitian

Hasil penelitian ini diharapkan mampu memberikan informasi kepada Pemerintah Provinsi Jawa Timur mengenai daerah-daerah di Provinsi Jawa Timur yang membutuhkan perhatian lebih dari pemerintah. Sehingga dalam rangka menghadapi era bonus demografi di kemudian hari, pemerintah daerah telah siap dengan kualitas SDM yang mumpuni. Selain itu, dengan adanya informasi tersebut diharapkan dapat membantu pemerintah daerah, baik pemerintah provinsi maupun kabupaten/kota dalam menentukan kebijakan yang tepat dengan memaksimalkan pembangunan SDM dan pembangunan ekonomi untuk menghadapi bonus demografi. Bagi pembaca, diharapkan hasil analisis ini bisa menjadi tambahan informasi terkait kondisi demografi Indonesia dan pembangunan nasional.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah variabel pembangunan SDM yang digunakan merupakan indikator pembangunan SDM dimensi kualitas berdasarkan Rencana Pembangunan Jangka Panjang Nasional 2005-2025. Sedangkan variabel pembangunan ekonomi yang digunakan disesuaikan dengan Anggaran Pendapatan dan Belanja Negara (APBN) tahun 2016. Selain itu, metode yang digunakan untuk pengelompokan data numerik adalah metode kluster hierarki *agglomerative* (*Single Linkage Method*, *Complete Linkage Method*, *Average Linkage*, dan *Ward's Method*) dengan jarak *Euclidean*. Metode pengelompokan

data kategori yang digunakan adalah metode ROCK. Kemudian hasil klaster dari kedua kelompok data tersebut digabungkan menggunakan metode *ensemble* ROCK.

BAB II

TINJAUAN PUSTAKA

2.1 Data Kategori dan Data Numerik

Data adalah suatu bentuk pencatatan berulang mengenai karakteristik suatu objek (Saefudin dkk, 2009). Menurut Kamus Besar Bahasa Indonesia (KBBI), data merupakan keterangan yang benar dan nyata. Sedangkan menurut Anderson, Sweeney, dan Williams (2012), data adalah fakta dan gambaran yang dikumpulkan, dianalisis, dan dirangkum untuk diinterpretasikan dalam presentasi. Berdasarkan jenis variabelnya, data dapat dibedakan menjadi data kategori dan data kuantitatif/ numerik. Data yang dapat dikelompokkan berdasarkan kategori spesifikasi tertentu disebut sebagai data kategori. Data kategori biasanya berisi data dengan skala nominal atau ordinal. Sedangkan data numerik adalah data yang menggunakan nilai sebenarnya sebagai ukuran dan mengandung skala pengukuran interval atau rasio.

2.2 Analisis Multivariat

Analisis multivariat merupakan analisis yang berkaitan dengan jumlah variabel lebih dari dua yang dianalisis secara simultan pada masing-masing pengamatan (Johnson & Wichern, 2007). Berikut ini merupakan struktur data dari pengamatan secara multivariat.

Tabel 2.1 Struktur Data Multivariat

Pengamatan ke-i	Variabel 1 (X_1)	Variabel 2 (X_2)	...	Variabel m (X_m)
1	$x_{1,1}$	$x_{1,2}$...	$x_{1,m}$
2	$x_{2,1}$	$x_{2,2}$...	$x_{2,m}$
.
.
.
n	$x_{n,1}$	$x_{n,2}$...	$x_{n,m}$

Terdapat beberapa teknik analisis multivariat yang sering digunakan, antara lain *Principal Component Analysis* (PCA) dan Analisis Faktor, *Multiple Regression*, Analisis Diskriminan, Analisis Korelasi Kanonikal, Analisis Multivariat dari varian dan kovarian, Analisis Klaster, Analisis Korespondensi, Analisis Gabungan, *Multidimensional Scalling*, dan *Structural Equation Modelling* (SEM) (Hair dkk, 2010).

2.3 Analisis Klaster

Analisis klaster atau pengelompokan berbeda dengan metode klasifikasi. Klasifikasi berkaitan dengan sejumlah kelompok yang telah diketahui sebelumnya dan tujuan operasionalnya adalah untuk menetapkan pengamatan baru. Analisis klaster adalah teknik yang lebih primitif karena tidak ada asumsi yang dibuat mengenai jumlah kelompok atau struktur kelompok sebelumnya. Pengelompokan dilakukan atas dasar kemiripan atau jarak (ketidak-miripan). Analisis klaster mengelompokkan objek-objek yang paling dekat kesamaannya dengan objek lain dan mempunyai kemiripan satu dengan yang lain dalam kelompok yang sama (Johnson & Wichern, 2007).

Salah satu faktor yang sangat berpengaruh terhadap hasil dari kelompok yang dibentuk adalah jarak antar objek pengamatan (Sharma, 1996). Oleh karena itu, dibutuhkan suatu alat ukur untuk menentukan jarak antar objek pengamatan. Berikut ini merupakan metode-metode pengukuran jarak antar objek ke- i (x_i) dengan objek ke- j (x_j) berdasarkan karakteristik variabel yang dikelompokkan.

a. Metode pengukuran jarak untuk variabel kategori biner

Bila variabel yang diamati berupa variabel biner yang hanya memiliki dua macam kategori yang berbeda (0,1), maka variabel yang diamati dapat dibentuk suatu tabel kontingensi seperti ditunjukkan pada Tabel 2.2. Perhitungan ukuran jarak antara variabel x_i dan x_j untuk pengukuran data biner dapat menggunakan ukuran yang disajikan pada Tabel 2.3.

Tabel 2.2 Tabel Kontingensi Data Biner

Kategori x_i	Kategori x_j		Total
	1	0	
1	a	b	a + b
0	c	d	c + d
Total	a + c	b + d	a + b + c + d

Tabel 2.3 Ukuran Jarak Data Biner

Jenis	Rumus
<i>Russel and Rao</i>	$RR(x_i, x_j) = \frac{a}{a + b + c + d}$
<i>Simple Matching</i>	$SM(x_i, x_j) = \frac{a + d}{a + b + c + d}$
<i>Jaccard</i>	$JACCARD(x_i, x_j) = \frac{a}{a + b + c}$
<i>Dice Czekanowski, Sørensen</i>	$DICE(x_i, x_j) = \frac{2a}{2a + b + c}$

- b. Metode pengukuran jarak untuk variabel kategori nominal
 Pada pengamatan dengan variabel nominal maka pengukuran memiliki konsep yang sama dengan *simple matcing coefficient* maupun *dice*, dimana kategorinya dapat lebih dua macam. Dengan jumlah variabel sebanyak m , maka rumus untuk pengukuran jarak variabel nominal antara x_i dan x_j ditunjukkan oleh persamaan berikut.

$$sim(x_i, x_j) = \frac{1}{m} \sum_{l=1}^m S_{ijl} \quad (2.1)$$

Dimana $S_{ijl} = 1$ jika $x_{il} = x_{jl}$ dan $S_{ijl} = 0$ jika $x_{il} \neq x_{jl}$.

- c. Metode pengukuran jarak untuk variabel kategori ordinal
 Pada pengamatan dengan variabel ordinal, pengukuran yang digunakan memiliki konsep yang sama dengan metode untuk data numerik, dimana kategorinya dinyatakan sebagai suatu bilangan bulat. Salah satu metode yang dapat digunakan untuk variabel ordinal adalah jarak *manhattan*. Dengan jumlah variabel sebanyak m , maka rumus untuk pengukuran jarak x_i

dan x_j pada variabel nominal ditunjukkan oleh persamaan berikut.

$$sim(x_i, x_j) = \sum_{l=1}^m |x_{il} - x_{jl}| \quad (2.2)$$

d. Metode pengukuran jarak untuk variabel numerik

Pada variabel yang memiliki jenis data numerik, maka ukuran jarak yang digunakan adalah jarak *Euclidean*. Misalkan terdapat dua observasi dengan variabel-variabel berdimensi m , yaitu $\mathbf{x}_i = [x_1, x_2, \dots, x_m]^T$ dan $\mathbf{x}_j = [x_1, x_2, \dots, x_m]^T$. Konsep jarak *Euclidean* yang mengukur jarak antara observasi \mathbf{x}_i dan \mathbf{x}_j adalah sebagai berikut.

$$d_{ij} = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)} \quad (2.3)$$

Terkait dengan pengertian dan tujuan dilakukannya analisis kluster, dapat dinyatakan bahwa suatu kelompok yang baik adalah kelompok yang memiliki ciri-ciri sebagai berikut.

- a. Homogenitas (kesamaan) yang tinggi antar anggota dalam satu kelompok (*within- cluster*).
- b. Heterogenitas (perbedaan) yang tinggi antar kelompok yang satu dengan kelompok yang lain (*between cluster*).

Tahap pengelompokan dalam analisis kluster dibedakan berdasarkan jenis data yang dimiliki. Pada umumnya analisis kelompok terfokus pada data numerik. Namun, terdapat juga kasus pengelompokan pada data kategori, atau bahkan data campuran numerik dan kategori. Analisis kluster pada data kategori tidak dapat dilakukan seperti pada data numerik. Hal tersebut dikarenakan data kategori memiliki sifat khusus, sehingga pengelompokan data kategori menjadi lebih rumit dibandingkan pengelompokan data numerik (Hair dkk, 2010).

2.4 Analisis Kluster Data Numerik

Analisis kluster untuk data numerik dilakukan berdasarkan ukuran ketidakmiripan atau jarak untuk data numerik. Hasil pengelompokan disajikan dalam bentuk *dendrogram* (diagram pohon) yang memungkinkan penelusuran objek-objek yang

diamati menjadi lebih mudah dan informatif. Teknik yang digunakan untuk pengelompokan meliputi metode hierarki dan metode non hierarki.

2.4.1 Klaster Metode Hierarki

Terdapat dua teknik pengelompokan dalam analisis klaster hierarki, yaitu teknik pembagian (*divisive*) dan teknik penggabungan (*agglomerative*). Pada penelitian ini, metode klaster yang digunakan adalah metode pengelompokan hierarki *agglomerative*. Metode-metode pengelompokan hierarki *agglomerative* dibedakan berdasarkan konsep jarak antar kelompok. Berikut merupakan penjelasan singkat mengenai macam-macam metode pada teknik *agglomerative*.

a. *Single Linkage*

Pengukuran jarak pada metode *single linkage* berdasarkan jarak terkecil atau terdekat. Jika kedua objek dipisahkan oleh jarak terpendek, maka dua objek tersebut dalam satu kelompok. Pengukuran jarak metode ini dapat ditulis dalam rumus berikut.

$$d_{w(i,j)} = \min \{d_{wi}, d_{wj}\} \quad (2.4)$$

Dimana $d_{w(i,j)}$ merupakan ukuran kemiripan antara kelompok ke- w dengan kelompok (i,j) yang merupakan penggabungan antara kelompok ke- i dan ke- j .

b. *Complete Linkage*

Pengukuran jarak pada metode *complete linkage* dilakukan berdasarkan jarak maksimum atau terjauh. Rumus yang digunakan dalam menentukan jarak pada metode ini adalah sebagai berikut.

$$d_{w(i,j)} = \max \{d_{wi}, d_{wj}\} \quad (2.5)$$

Dimana $d_{w(i,j)}$ merupakan ukuran kemiripan antara kelompok ke- w dengan kelompok (i,j) yang merupakan penggabungan antara kelompok ke- i dan ke- j .

c. *Average Linkage*

Kriteria yang digunakan dalam mengukur jarak pada metode ini adalah rata-rata jarak seluruh individu dalam suatu kelompok dengan jarak seluruh individu dalam kelompok lain. Dapat dituliskan dalam rumus sebagai berikut.

$$d_{w(i,j)} = \frac{\sum_q \sum_r d_{qr}}{N_{(ij)} N_w} \quad (2.6)$$

Dimana $d_{w(i,j)}$ merupakan ukuran kemiripan antara kelompok ke- w dengan kelompok (i,j) yang merupakan penggabungan antara kelompok ke- i dan ke- j . $N_{(ij)}$ dan N_w merupakan jumlah pengamatan dalam kelompok ij dan w . Sedangkan, d_{qr} merupakan jarak antara pengamatan ke- q dalam kelompok ij dengan pengamatan ke- r dalam kelompok w .

d. *Ward's Method*

Sedangkan *Ward's Method* menurut Johnson dan Wichern (2007) merupakan metode yang mencoba meminimumkan varian dalam kelompok. Jika klaster sebanyak K maka ESS sebagai jumlahan dari ESS_k atau $ESS = ESS_1 + ESS_2 + \dots + ESS_k$. Sehingga rumus untuk menghitung jarak antara dua klaster menggunakan *Ward's Method* adalah sebagai berikut.

$$ESS = \sum_{j=1}^n (x_j - \bar{x})'(x_j - \bar{x}) \quad (2.7)$$

Dimana x_j adalah nilai dari objek ke- j , dengan $j=1,2,\dots,n$ dan \bar{x} adalah rata-rata dari objek pengamatan.

Kelebihan dari metode hierarki, yaitu mempercepat pengolahan data karena hasil yang terbentuk berupa tingkatan, sehingga lebih mudah dipahami. Namun, metode ini sangat lemah pada data yang memiliki nilai *outlier*, perbedaan ukuran jarak yang digunakan dan variabel yang tidak sesuai.

2.4.2 Klaster Metode Non Hierarki

Berbeda dengan metode hierarki, metode non hierarki dimulai dengan terlebih dahulu jumlah kelompok yang diinginkan (2 kelompok, 3 kelompok, atau lebih). Setelah jumlah kelompok diketahui, selanjutnya dilakukan proses pengelompokan. Salah satu contoh metode non hierarki adalah *K-Means Cluster* (Hair dkk, 2010).

Prosedur yang perlu dilakukan dalam menggunakan metode *K-Means* adalah sebagai berikut (Johnson & Wichern, 2007).

1. Mempartisi atau mengelompokkan data ke dalam jumlah kluster yang telah ditentukan.
2. Menentukan jarak setiap data ke setiap pusatnya.
3. Mengelompokkan setiap data ke pusat terdekat.
4. Menghitung nilai pusat baru sebagai rata-rata data dalam kluster.
5. Melakukan prosedur di atas secara iteratif hingga tidak terdapat lagi perubahan dalam melakukan pengelompokan.

Jika objek dipindahkan dari konfigurasi awal maka nilai pusat (*mean*) perlu diperbarui sebelum melanjutkan ke langkah berikutnya. Jika objek ke- j masuk ke dalam kluster maka menggunakan rumus berikut.

$$\bar{x}_{i,baru} = \frac{n\bar{x}_i + x_{ji}}{n + 1} \quad (2.8)$$

Sedangkan jika objek ke- j keluar dari kluster maka rumus yang digunakan adalah sebagai berikut.

$$\bar{x}_{i,baru} = \frac{n\bar{x}_i + x_{ji}}{n - 1} \quad (2.9)$$

Dimana n adalah jumlah objek dari kluster sebelumnya dengan *centroid* $\bar{x}' = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p)$ (Johnson & Wichern, 2007).

Jika dalam pengelompokan objek tidak terdapat informasi apriori tentang kelompok yang terbentuk, maka metode yang tepat untuk digunakan adalah metode hierarki. Dengan kata lain hasil dari metode hierarki dapat digunakan sebagai informasi awal dalam mengelompokkan objek menggunakan metode non hierarki

2.5 Analisis Kluster Data Kategori

Analisis kluster untuk data kategori dilakukan menggunakan ukuran kemiripan atau jarak untuk data kategori. Kemudian dapat dilakukan pengelompokan menggunakan metode hierarki maupun non hierarki. Namun, metode hierarki dan non hierarki tradisional dinilai tidak cocok digunakan pada data kategori. Oleh karena itu, telah dikembangkan metode yang sesuai, yaitu metode

ROCK (*RObust Clustering with linKs*) untuk mengelompokkan data kategori tersebut (Guha, Rastogi, & Shim, 2000).

Metode ROCK dikembangkan dari metode pengelompokan hierarki *agglomerative* yang digunakan untuk data kategori. Algoritma pengelompokan hierarki tradisional yang menggunakan jarak antar titik untuk pengelompokan dinilai kurang sesuai untuk data atribut dan kategori. Oleh karena itu, dibentuk sebuah konsep baru, yaitu *link* untuk mengukur kesamaan/ kedekatan antar sepasang titik data. Pengamatan yang mempunyai tingkat hubungan (*link*) tinggi digabungkan ke dalam satu kelompok, sedangkan yang mempunyai tingkat hubungan (*link*) yang kecil dipisahkan dari kelompok dimana data tersebut dikelompokkan. Metode ROCK merupakan metode yang dapat menangani *outlier* dengan cukup efektif. Pemangkasan *outlier* memungkinkan untuk membuang data yang tidak memiliki tetangga, sehingga titik tersebut tidak berpartisipasi dalam pengelompokan. Namun, dalam beberapa situasi, *outlier* dapat hadir sebagai kelompok-kelompok kecil (Guha, Rastogi, & Shim, 2000).

Pengelompokan data kategori dengan algoritma ROCK dilakukan dengan tiga langkah. Langkah pertama adalah menghitung kesamaan. Ukuran kemiripan antar pasangan objek ke- i dan objek ke- j dihitung dengan rumusan yang didefinisikan sebagai berikut.

$$\text{sim}(X_i, X_j) = \frac{|X_i \cap X_j|}{|X_i \cup X_j|}, i \neq j \quad (2.10)$$

Dengan,

$i = 1, 2, 3, \dots, n$ dan $j = 1, 2, 3, \dots, n$

X_i : Himpunan pengamatan ke- i , dengan $X_i = \{x_{1i}, x_{2i}, \dots, x_{m_{\text{kategori } i}}\}$,

X_j : Himpunan pengamatan ke- j dengan $X_j = \{x_{1j}, x_{2j}, \dots, x_{m_{\text{kategori } j}}\}$,

$|X|$: Bilangan kardinal atau jumlah anggota dari himpunan X .

Langkah kedua adalah menentukan tetangga, pengamatan dinyatakan sebagai tetangga jika nilai $\text{sim}(X_i, X_j) \geq \theta$. Langkah

terakhir adalah menghitung *link* antar objek pengamatan. Besarnya *link* dipengaruhi oleh nilai *threshold* (θ) yang merupakan paramater yang ditentukan oleh pengguna yang dapat digunakan untuk mengontrol seberapa dekat hubungan antar objek. Besarnya nilai θ yang diinputkan adalah $0 < \theta < 1$. Algoritma ROCK berhenti ketika jumlah dari kelompok yang diharapkan sudah terpenuhi atau tidak ada lagi *link* antara kelompok-kelompok (Dutta, Mahanta, & Arun, 2005).

Metode ROCK menggunakan informasi tentang *link* sebagai ukuran kemiripan antar objek. Jika terdapat objek pengamatan X_i , X_j , dan X_k , dimana X_i merupakan tetangga dari X_j dan X_j merupakan tetangga dari X_k . Oleh karena itu, dapat dikatakan bahwa X_i memiliki *link* dengan X_k walaupun X_i bukan tetangga dari X_k . Cara menghitung jumlah *link* untuk semua kemungkinan pasangan dari n objek adalah menggunakan matriks **A**. Matriks **A** merupakan matriks berukuran $n \times n$ yang bernilai 1, jika X_i dan X_j dinyatakan mirip (tetangga) dan bernilai 0, jika X_i dan X_j dinyatakan tidak mirip (bukan tetangga). Jumlah *link* antar pasangan X_i dan X_j diperoleh dari hasil kali antara baris ke X_i dan kolom ke X_j dari matriks **A**. Jika *link* antara X_i dan X_j semakin besar, maka semakin besar kemungkinan X_i dan X_j berada dalam satu kelompok yang sama (Dutta, Mahanta, & Arun, 2005).

Penggabungan kelompok menggunakan algoritma ROCK didasarkan atas ukuran kebaikan (*goodness measure*) antar kelompok dengan rumus seperti pada persamaan 2.11. *Goodness measure* adalah persamaan untuk menghitung jumlah *link* dibagi dengan kemungkinan *link* yang terbentuk berdasarkan ukuran kelompoknya (Tyagi & Sharma, 2012).

$$g(C_i, C_j) = \frac{\text{link}[C_i, C_j]}{(n_i + n_j)^{1+2f(\theta)} - n_i^{1+2f(\theta)} - n_j^{1+2f(\theta)}} \quad (2.11)$$

Dengan $\text{link}[C_i, C_j] = \sum_{X_i \in C_i, X_j \in C_j} \text{link}(X_i, X_j)$ yang menyatakan

bahwa jumlah *link* dari semua kemungkinan pasangan objek yang ada dalam C_i dan C_j serta n_i dan n_j , masing-masing menyatakan

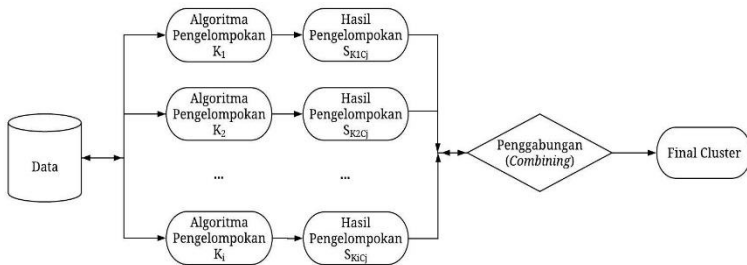
jumlah anggota dalam kelompok ke- i dan j , sedangkan

$$f(\theta) = \frac{1 - \theta}{1 + \theta}.$$

2.6 Analisis Klaster Data Campuran

Analisis klaster terhadap data campuran diawali dengan membagi data menjadi dua, yaitu data murni numerik dan data murni kategori. Apabila terdapat data dengan variabel berskala campuran sebanyak m , dengan m_{numerik} merupakan jumlah variabel numerik, dan m_{kategori} merupakan jumlah variabel kategori. Maka dari keterangan tersebut diperoleh $m = m_{\text{numerik}} + m_{\text{kategori}}$. Selanjutnya dilakukan pengelompokan data sesuai dengan jenis data secara terpisah. Hasil pengelompokan tersebut kemudian digabungkan menggunakan metode pengelompokan *ensemble* berdasarkan ROCK sehingga diperoleh kelompok akhir (*final cluster*).

Pengelompokan *ensemble* merupakan metode yang menggabungkan algoritma yang berbeda untuk mendapatkan partisi umum dari data dengan tujuan untuk konsolidasi dari portofolio hasil pengelompokan individu (Suguna & Selvi, 2012). Tujuan pengelompokan *ensemble* adalah untuk menggabungkan hasil pengelompokan dari beberapa algoritma pengelompokan untuk mendapatkan hasil pengelompokan yang lebih baik dan *robust* (Yoon, Ahn, Lee, Cho, & Kim, 2006). Pengelompokan *ensemble* terdiri dari dua tahap algoritma. Tahap pertama adalah melakukan pengelompokan dengan beberapa algoritma dan menyimpang hasil pengelompokan tersebut. Kedua, menggunakan fungsi konsensus untuk menentukan *final cluster* dari kelompok-kelompok hasil tahap pertama. Skema pengelompokan *ensemble* secara umum ditunjukkan pada Gambar 2.1.

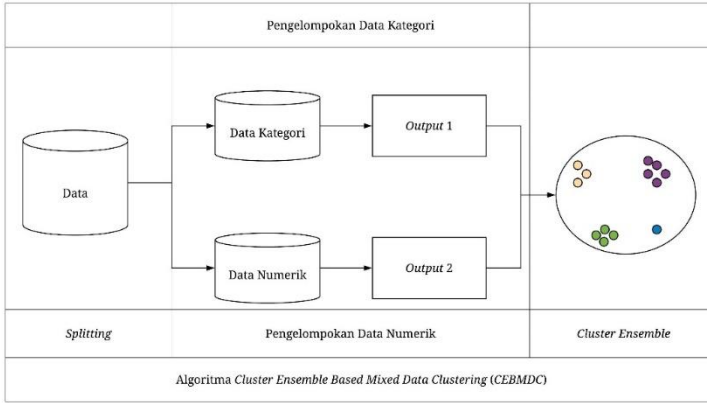


Sumber : Irfandi (2016)

Gambar 2.1 Tahapan Umum Pengelompokan *Ensemble*

Dalam penelitian ini pengelompokan data campuran dilakukan menggunakan Algoritma CEBMDC yang secara umum ditunjukkan oleh Gambar 2.2. Berikut ini adalah langkah-langkah dalam analisis data campuran menggunakan Algoritma CEBMDC (He, Xu, & Deng, 2005).

- Membagi data menjadi dua bagian, yaitu data murni numerik dan murni kategori.
- Melakukan pengelompokan objek dengan variabel numerik dengan algoritma pengelompokan data numerik, serta melakukan pengelompokan objek dengan variabel kategori dengan algoritma pengelompokan data kategori.
- Menggabungkan (*combining*) hasil pengelompokan dari variabel numerik dan kategori yang disebut dengan proses *ensemble*.
- Melakukan pengelompokan *ensemble* menggunakan algoritma pengelompokan data kategori untuk mendapatkan kelompok akhir (*final cluster*).



Sumber : Irfandi (2016)

Gambar 2.2 Tahapan Umum Algoritma CEBMDC

2.7 Kinerja Hasil Pengelompokan

Pengukuran kinerja hasil pengelompokan merupakan langkah untuk mengetahui validitas suatu pengelompokan. Kelompok yang baik adalah kelompok yang memiliki kehomogenan yang tinggi antar anggota kelompok dan keheterogenan yang tinggi antar kelompok (Hair dkk, 2010).

2.7.1 Kinerja Hasil Pengelompokan Data Numerik

Indeks *R-Squared* merupakan satu indeks yang dapat digunakan untuk menentukan jumlah kelompok optimum pada pengelompokan hierarki (Sharma, 1996). Indeks tersebut melibatkan perhitungan keragaman data, baik keragaman total, keragaman dalam kelompok, maupun keragaman data antar kelompok. Indeks validitas untuk menentukan jumlah kelompok optimum pada pengelompokan hierarki dapat dituliskan sebagai berikut.

Sum of Square Total (SST) :

$$SST = \sum_{l=1}^{m_{numerik}} \sum_{i=1}^n (x_{il} - \bar{x}_l)^2 \quad (2.12)$$

Sum of Square Within Group (SSW) :

$$SSW = \sum_{c=1}^C \sum_{l=1}^{m_{\text{numerik}}} \sum_{i=1}^n (x_{ilc} - \bar{x}_{lc})^2 \quad (2.13)$$

Sum of Square Between Group (SSB) :

$$SSB = SST - SSW \quad (2.14)$$

dengan keterangan sebagai berikut.

m_{numerik} : Jumlah variabel numerik dalam pengamatan

C : Jumlah kelompok yang dibentuk dalam pengamatan

n : Jumlah objek pengamatan

n_c : Jumlah anggota pada kelompok ke- c , untuk $c = 1, 2, \dots, C$

\bar{x}_l : Rata-rata keseluruhan objek pada variabel ke- l , untuk

$l = 1, 2, \dots, m_{\text{numerik}}$

\bar{x}_{lc} : Rata-rata variabel ke- l pada kelompok ke- c , untuk

$c = 1, 2, \dots, C$.

R^2 kelompok baru adalah rasio dari SSB dan SST , artinya R^2 dapat didefinisikan sebagai ukuran perbedaan antar kelompok dengan nilai berkisaran antara 0 sampai 1. Nilai R^2 sama dengan 0 menunjukkan bahwa tidak terdapat perbedaan antar kelompok. Sementara R^2 sama dengan satu menunjukkan adanya perbedaan yang signifikan antar kelompok yang ter-bentuk. Berikut merupakan rumus perhitungan R^2 .

$$R^2 = \frac{SSB}{SST} = \frac{[SST - SSW]}{SST} \quad (2.15)$$

Penentuan jumlah kelompok optimum yang terbentuk dapat dilihat dari plot indeks R^2 . Jika terdapat kenaikan nilai indeks R^2 yang sangat tajam pada suatu kelompok, maka pada titik tersebut merupakan jumlah kelompok optimum yang terbentuk. Selain menggunakan nilai R^2 , penentuan jumlah kelompok optimum juga dapat dilakukan dengan melihat nilai maksimum dari *Pseudo-F*. Berikut merupakan rumus yang digunakan untuk menghitung nilai *Pseudo-F*.

$$Pseudo - F = \frac{\left(\frac{R^2}{c-1} \right)}{\left(\frac{1-R^2}{n-c} \right)} \quad (2.16)$$

Setelah mendapatkan jumlah kelompok yang optimum dalam setiap metode. Selanjutnya, proses analisis dilanjutkan dengan menentukan metode pengelompokan yang paling optimum berdasarkan nilai *icdrate* minimum. Berikut ini merupakan rumus perhitungan nilai *icdrate*.

$$icdrate = 1 - R^2 \quad (2.17)$$

Semakin kecil nilai *icdrate*, artinya semakin kecil perbedaan di dalam anggota kelompok tersebut.

2.7.2 Kinerja Hasil Pengelompokan Data Kategori

Ukuran keragaman untuk data kategori telah dikembangkan oleh Light dan Margolin (1971), Okada (Okada, 1999), serta Kader dan Perry (2007). Jika terdapat sebanyak n pengamatan dengan n_k merupakan jumlah pengamatan dengan kategori ke- k dimana $k =$

1,2,3, ..., K dan $\sum_{k=1}^K n_k = n$. Selanjutnya, n_{kc} merupakan jumlah

pengamatan dengan kategori ke- k dan kelompok ke- c , dimana $c =$ 1,2,3, ..., C dengan C adalah jumlah kelompok yang terbentuk.

Sehingga $n_k = \sum_{c=1}^C n_{kc}$ merupakan jumlah pengamatan pada

kategori ke- k . Total jumlah pengamatan dapat dituliskan menjadi

$$n = \sum_{c=1}^C n_c = \sum_{k=1}^K n_k = \sum_{k=1}^K \sum_{c=1}^C n_{kc}.$$

Jumlah kuadrat total (*SST*) untuk sebuah variabel dengan data kategorik dapat dirumuskan seperti persamaan (2.18). Untuk total jumlah kuadrat dalam kelompok (*SSW*) dirumuskan dalam persamaan (2.19). Serta jumlah kuadrat antar kelompok (*SSB*) dapat dirumuskan seperti persamaan (2.20) (Dewi, 2012).

$$SST = \frac{n}{2} - \frac{1}{2n} \sum_{k=1}^K n_k^2 \quad (2.18)$$

$$SSW = \sum_{c=1}^C \left(\frac{n_c}{2} - \frac{1}{2n_c} \sum_{k=1}^K n_{kc}^2 \right) = \frac{n}{2} - \frac{1}{2} \sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \quad (2.19)$$

$$SSB = \frac{1}{2} \left(\sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2 \quad (2.20)$$

Sedangkan *mean of squares total (MST)*, *mean of squares within (MSW)*, dan *mean of squares between (MSB)* dapat dirumuskan sebagai berikut.

$$MSS = \frac{SST}{(n-1)} \quad (2.21)$$

$$MSW = \frac{SSW}{(n-C)} \quad (2.22)$$

$$MSB = \frac{SSB}{C-1} \quad (2.23)$$

Simpangan baku dalam kelompok (S_w) dan simpangan baku antar kelompok (S_B) untuk data kategori dapat dirumuskan sebagai berikut.

$$S_w = [MSW]^{1/2} \quad (2.24)$$

$$S_B = [MSB]^{1/2} \quad (2.25)$$

Seperti hanya dengan data numerik, kinerja suatu metode pengelompokan untuk data kategori semakin baik apa bila nilai rasio S_w dan S_B semakin kecil. Artinya terdapat homogenitas maksimum dalam kelompok dan heterogenitas maksimum antar kelompok (Bunkers & James, 1996).

2.8 Pembangunan Nasional

Pada penjelasan Undang-Undang Republik Indonesia nomor 17 tahun 2007 tentang Rencana Pembangunan Jangka Panjang Nasional 2005-2025 menyatakan bahwa, pembangunan nasional

adalah rangkaian upaya pembangunan yang berkesinambungan yang meliputi seluruh aspek kehidupan masyarakat, bangsa, dan negara untuk melaksanakan tugas mewujudkan tujuan nasional sebagaimana dirumuskan dalam Pembukaan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945. Rangkaian upaya pembangunan tersebut memuat kegiatan-kegiatan pembangunan yang berlangsung tanpa henti, dengan menaikkan tingkat kesejahteraan masyarakat dari generasi demi generasi. Pelaksanaan upaya tersebut dilakukan dalam konteks memenuhi kebutuhan masa sekarang tanpa mengurangi kemampuan generasi yang akan datang untuk memenuhi kebutuhannya. Dalam visi dan arah Pembangunan Jangka Panjang (PJP) Tahun 2005-2025 dijelaskan bahwa pembangunan nasional terdiri dari pembangunan politik, pembangunan pertahanan dan keamanan, pembangunan hukum dan penyelenggaraan negara, pembangunan sosial budaya, pembangunan SDM, pembangunan ekonomi, pembangunan daerah, pembangunan infrastruktur, serta pembangunan SDA dan lingkungan hidup.

2.8.1 Pembangunan SDM

Pembangunan sumber daya manusia (SDM) pada dasarnya merupakan pembangunan manusia sebagai subyek (*human capital*), obyek (*human resources*), dan penikmat pembangunan yang mencakup seluruh siklus hidup manusia sejak dalam kandungan sampai dengan akhir hidupnya. Manusia, untuk selanjutnya disebut sebagai penduduk dan sesuai kepentingannya dibedakan atas jenis kelamin (laki-laki dan perempuan), kelompok umur (anak, remaja, pemuda, usia produktif, usia reproduktif, dan usia lanjut), serta kelompok miskin dan rentan. Dimensi pembangunan SDM dapat dilihat dari tiga aspek utama, yaitu kualitas, kuantitas, dan mobilitas penduduk. Kualitas penduduk tercermin dari tingkat kesejahteraan penduduk, yaitu tingkat kesehatan dan gizi, pendidikan, produktivitas, dan akhlak mulia, menuju kepada pencapaian kesejahteraan sosial yang baik. Secara umum kualitas penduduk dapat dilihat dari peringkat atau nilai Indeks Pembangunan Manusia (IPM). Sedangkan, kuantitas

penduduk dikaitkan dengan jumlah dan laju pertumbuhannya. Aspek mobilitas penduduk merupakan refleksi dari perpindahan dan persebaran penduduk, yang merespon pembangunan ekonomi wilayah.

2.8.2 Pembangunan Ekonomi

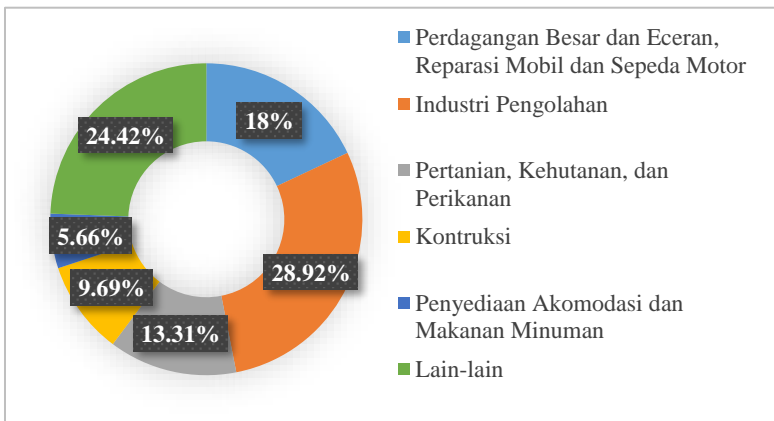
Menurut Sadono Sukirno (2012) pertumbuhan dan pembangunan ekonomi memiliki definisi yang berbeda. Pertumbuhan ekonomi adalah proses kenaikan *output* perkapita yang terus menerus dalam jangka panjang. Pertumbuhan ekonomi tersebut merupakan salah satu indikator keberhasilan pembangunan ekonomi. Oleh karena itu, semakin tinggi pertumbuhan ekonomi maka semakin tinggi pula kesejahteraan penduduk. Meskipun terdapat indikator yang lain, yaitu distribusi pendapatan. Sedangkan, pembangunan ekonomi adalah usaha meningkatkan pendapatan perkapita dengan jalan mengolah kekuatan ekonomi potensial menjadi ekonomi riil melalui penanaman modal, penggunaan teknologi, penambahan pengetahuan, peningkatan keterampilan, penambahan kemampuan berorganisasi, dan manajemen.

2.9 Gambaran Umum Provinsi Jawa Timur

Provinsi Jawa Timur merupakan provinsi dengan luas wilayah 47799.75 km² dan terdiri dari 38 kabupaten/kota. Jumlah penduduk Provinsi Jawa Timur adalah sebesar 39.1 juta jiwa dan merupakan jumlah penduduk terbesar kedua di Indonesia. Sedangkan laju pertumbuhan penduduknya pada tahun 2016 tercatat sekitar 0.59 persen per tahun dan TPT (Tingkat Pengangguran Terbuka) sebesar 4.21 persen. IPM (Indeks Pembangunan Manusia) di Jawa Timur pada tahun 2016 memiliki nilai sebesar 69.74. Meskipun mengalami kenaikan dari tahun sebelumnya, tapi angka tersebut masih belum mencapai target APBN (Anggaran Pendapatan dan Belanja Negara) tahun 2016 yang sebesar 70.1. Masalah kesenjangan ekonomi dan kemiskinan masih menjadi pekerjaan rumah yang cukup berat bagi Provinsi Jawa Timur. Berdasarkan data BPS pada tahun 2016, indeks gini Jawa Timur sebesar 0.40. Artinya tingkat ketimpangan yang ada di

Provinsi Jawa Timur tergolong sedang. Indeks gini sebesar 0.40 tersebut cukup besar jika dibandingkan pada tahun 2010 yang hanya sebesar 0.31. Namun, jika dibandingkan dengan tahun 2015, indeks gini Jawa Timur mengalami penurunan sebesar 0.02. Angka tersebut tidak cukup signifikan jika melihat pertumbuhan ekonomi di Jawa Timur yang mencapai 5.57 persen hingga kuartal III tahun 2016. Pertumbuhan ekonomi mengalami kenaikan 0.23 persen dibandingkan dengan tahun sebelumnya.

Struktur perekonomian Jawa Timur menurut lapangan usaha tahun 2016 didominasi oleh tiga lapangan usaha utama, yaitu kategori industri pengolahan (28.92 persen), kategori perdagangan besar dan eceran, reparasi mobil dan sepeda motor (18 persen), dan kategori pertanian, kehutanan, dan perikanan (13.31 persen). Berikut ini merupakan struktur PDRB (Produk Domestik Regional Bruto) Jawa Timur menurut lapangan usaha tahun 2016.



Sumber : BAPPEDA (2017)

Gambar 2.3 Struktur PDRB Jawa Timur Tahun 2016

2.10 Penelitian Sebelumnya

Pada subbab latar belakang telah dijelaskan mengenai berbagai macam metode klaster yang digunakan pada beberapa penelitian sebelumnya dan alasan pemilihan metode *ensemble* ROCK pada penelitian ini. Sedangkan untuk proses pemilihan

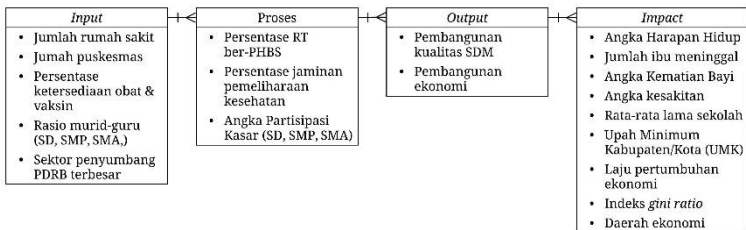
variabel penelitian, peneliti menggunakan acuan penelitian yang telah dilakukan oleh Dwi Maumere Putra dan Meita Nike Harmanto.

Putra (2015) melakukan penelitian dengan judul “Pemodelan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur Dengan Menggunakan Metode Regresi Logistik Ridge”. Variabel Y dalam penelitian tersebut adalah nilai IPM dari masing-masing kabupaten/kota di Jawa Timur. Sedangkan variabel X yang digunakan merupakan hasil dari indikator IPM, yaitu variabel kesehatan, pendidikan, dan ekonomi. variabel kesehatan diwakili oleh indikator Angka Kematian Bayi (AKB), presentase keluhan kesehatan, dan jumlah sarana kesehatan. Variabel pendidikan digambarkan oleh indikator angka buta huruf (Usia 10 tahun ke atas), Angka Partisipasi Sekolah (APS) SMA, rasio guru-murid (SMA), dan rasio sekolah-murid (SMA). Sedangkan variabel ekonomi digambarkan oleh indikator rumah tangga miskin, PDRB perkapita, pertumbuhan ekonomi, persentase penduduk usia 15 tahun ke atas yang bekerja, Tingkat Pengangguran Terbuka (TPT), dan Tingkat Partisipasi Angkatan Kerja (TPAK).

Pada tahun berikutnya, Harmanto (2016) melakukan analisis pengelompokan mengenai perubahan struktur kependudukan dalam menghadapi era bonus demografi di kabupaten/ kota Provinsi Jawa Timur. Metode yang digunakan dalam penelitian tersebut adalah analisis faktor, analisis klaster, analisis biplot, dan analisis diskriminan. Sedangkan variabel yang digunakan adalah kependudukan, ekonomi, kesejahteraan penduduk, dan ketenagakerjaan. Variabel kependudukan meliputi penduduk usia 0-14 tahun, penduduk usia 15-64 tahun, penduduk usia 65 tahun ke atas, kepadatan penduduk, *sex ratio*, migrasi *in*, dan migrasi *out*. Variabel ekonomi meliputi PDRB pertanian, PDRB manufaktur, PDRB jasa, dan UMR. Variabel kesejahteraan penduduk digambarkan oleh IPM, Angka Partisipasi Kasar (APK) SMP, APK SMA, Angka Partisipasi Menengah SMP, APM SMA, AKB, dan AKI. Serta variabel ketenagakerjaan yang meliputi TPAK, TPT, pekerja pertanian, pekerja manufaktur, dan pekerja jasa.

2.11 Kerangka Konsep Penelitian

Berdasarkan uraian sebelumnya telah didapatkan beberapa variabel yang akan digunakan dalam penelitian ini. Kualitas SDM yang tercermin dari nilai IPM digunakan sebagai acuan pemilihan variabel berdasarkan indikator pendidikan, kesehatan, dan standar hidup layak. Indikator pendidikan diwakili oleh variabel rata-rata lama sekolah, APK jenjang pendidikan SD, SMP, dan SMA, serta variabel rasio murid-guru jenjang pendidikan SD, SMP, dan SMA. Sedangkan indikator kesehatan diwakili oleh variabel jumlah rumah sakit, jumlah puskesmas, persentase ketersediaan obat dan vaksin, persentase Rumah Tangga (RT) berperilaku hidup bersih dan sehat, persentase penduduk yang memiliki jaminan pemeliharaan kesehatan, AHH, jumlah ibu yang meninggal, AKB, dan angka kesakitan. Sedangkan indikator standar hidup layak dan pembangunan ekonomi diwakili oleh variabel Upah Minimum Kabupaten/Kota (UMK), daerah ekonomi, laju pertumbuhan ekonomi, indeks *gini ratio*, dan sektor penyumbang PDRB terbesar. Beberapa variabel tersebut kemudian dipisahkan berdasarkan variabel *input*, *proses*, *output*, dan *impact*. Penjelasan lebih lanjut mengenai variabel-variabel yang digunakan pada penelitian ini dapat dilihat pada kerangka konsep berikut.



Gambar 2.4 Kerangka Konsep Penelitian

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Jawa Timur, hasil publikasi Badan Pusat Statistika (BPS) dan Badan Perencanaan Pembangunan Daerah Pemerintah Provinsi Jawa Timur (Bappeda), Peraturan Menteri Komunikasi dan Informatika Republik Indonesia Nomor 39 Tahun 2008, serta Peraturan Gubernur Jawa Timur Nomor 68 Tahun 2015 tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2016. Data tersebut mencakup indikator keberhasilan pembangunan nasional di bidang kualitas SDM dan ekonomi pada tahun 2016. Adapun unit penelitian yang digunakan berupa 38 kabupaten/kota di Provinsi Jawa Timur.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini terdiri dari variabel berskala numerik dan kategori. Terdapat 21 variabel dengan keterangan sebagai berikut.

Tabel 3.1 Variabel Penelitian Berskala Numerik

Variabel	Keterangan
X ₁	Jumlah rumah sakit
X ₂	Jumlah puskesmas
X ₃	Persentase ketersediaan obat dan vaksin
X ₄	Persentase RT berperilaku hidup bersih dan sehat
X ₅	Persentase kepemilikan jaminan pemeliharaan kesehatan
X ₆	Angka Harapan Hidup
X ₇	Jumlah ibu yang meninggal
X ₈	Angka Kematian Bayi
X ₉	Angka Kesakitan

Tabel 3.1 Variabel Penelitian Berskala Numerik (lanjutan)

Variabel	Keterangan
X ₁₀	Rasio murid-guru SD
X ₁₁	Rasio murid-guru SMP
X ₁₂	Rasio murid-guru SMA
X ₁₃	APK SD
X ₁₄	APK SMP
X ₁₅	APK SMA
X ₁₆	Rata-rata lama sekolah
X ₁₇	Upah Minimum Kabupaten/Kota (UMK)

Tabel 3.2 Variabel Penelitian Berskala Kategori

Variabel	Keterangan	Kategori
X ₁₈	Indeks <i>gini ratio</i>	0 : Belum mencapai target pembangunan nasional
		1 : Telah mencapai target pembangunan nasional
X ₁₉	Laju pertumbuhan ekonomi	0 : Belum mencapai target pembangunan nasional
		1 : Telah mencapai target pembangunan nasional
X ₂₀	Daerah ekonomi	0 : Zona 1
		1 : Zona 2
		2 : Zona 3
		3 : Zona 4
		4 : Zona 5
X ₂₁	Sektor penyumbang PDRB Terbesar	0 : Industri pengolahan
		1 : Perdagangan besar dan eceran, reparasi mobil dan sepeda motor
		2 : Pertanian, kehutanan, dan perikanan
		3 : Lainnya

Berdasarkan Tabel 3.1 dan Tabel 3.2, definisi operasional dari masing-masing variabel tersebut adalah sebagai berikut.

1. Jumlah fasilitas kesehatan, yaitu jumlah fasilitas pelayanan kesehatan yang digunakan untuk menyelenggarakan upaya pelayanan kesehatan perorangan, baik promotif, preventif, kuratif, maupun rehabilitatif yang diberikan oleh pemerintah, pemerintah daerah, atau masyarakat. Dalam penelitian ini, peneliti menggunakan fasilitas kesehatan berupa jumlah rumah sakit dan jumlah puskesmas.
2. Persentase ketersediaan obat dan vaksin adalah indikator yang digunakan pemerintah dalam mengukur capaian upaya pelayanan kesehatan (menjamin ketersediaan jenis obat-obatan) berdasarkan indikator program terbaru (20 jenis obat). Target yang ingin dicapai pemerintah pada tahun 2016 adalah 80%.
3. Persentase Rumah Tangga (RT) berperilaku hidup bersih dan sehat, yaitu persentase rumah tangga yang melaksanakan 10 indikator PHBS (Perilaku Hidup Bersih dan Sehat).
4. Persentase jaminan pemeliharaan kesehatan, yaitu persentase masyarakat yang mengikuti program jaminan kesehatan (askes, jaminan kesehatan nasional, BPJS kesehatan).
5. Angka Harapan Hidup (AHH), yaitu rata-rata tahun hidup yang masih akan dijalani oleh seseorang yang telah berhasil mencapai umur x, pada suatu tahun tertentu dalam situasi mortalitas yang berlaku di lingkungan masyarakatnya.
6. Jumlah ibu yang meninggal adalah banyaknya kematian perempuan pada saat hamil atau selama 42 hari sejak terminasi kehamilan tanpa memandang lama dan tempat persalinan yang disebabkan karena kehamilannya atau pengelolaannya dan bukan karena sebab-sebab lain. Berdasarkan data yang didapatkan, jumlah kematian ibu merupakan penjumlahan dari jumlah kematian ibu hamil, jumlah kematian ibu bersalin, dan jumlah kematian ibu nifas.
7. Angka Kematian Bayi (AKB), yaitu angka yang menunjukkan banyaknya kematian bayi usia 0 tahun dari setiap 1000 kelahiran hidup pada tahun tertentu. Dapat dikatakan juga

sebagai probabilitas bayi meninggal sebelum mencapai usia satu tahun.

8. Angka kesakitan, yaitu perbandingan antara jumlah penduduk yang sakit dengan jumlah penduduk pertengahan tahun.
9. Rasio guru-murid merupakan angka yang menunjukkan perbandingan antara jumlah murid dan guru pada suatu jenjang pendidikan (SD, SMP, SMA) dengan jumlah sekolah yang bersangkutan.
10. Angka Partisipasi Kasar (APK), yaitu proporsi anak sekolah pada suatu jenjang tertentu (SD, SMP, SMA) pada kelompok usia tertentu (7-12 tahun, 13-15 tahun, 16-18 tahun).
11. Rata-rata lama sekolah, yaitu jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal (tidak termasuk tahun yang mengulang).
12. Upah Minimum Kabupaten/Kota (UMK) adalah suatu standar minimum yang digunakan oleh para pengusaha atau pelaku industri untuk memberikan upah kepada pegawai, karyawan, atau buruh di dalam lingkungan usaha atau kerjanya. UMK berlaku di suatu wilayah kabupaten/kota berdasarkan peraturan yang telah ditetapkan oleh gubernur di provinsi tersebut.
13. Indeks *gini ratio* adalah ukuran pemerataan pendapatan yang dihitung berdasarkan kelas pendapatan. Nilai indeks gini berkisar antara 0 sampai 1. Indeks gini bernilai 0 menunjukkan adanya pemerataan pendapatan yang sempurna, atau setiap orang memiliki pendapatan yang sama. Pada penelitian ini indeks gini dari masing-masing kabupaten/kota dikategorikan berdasarkan pencapaian terhadap target pembangunan, yaitu 0.39. Sehingga akan dikategorikan 1 apabila memiliki indeks gini kurang dari sama dengan 0.39, dan akan dikategorikan 0 apabila memiliki indeks gini lebih dari 0.39.
14. Laju pertumbuhan ekonomi, yaitu perkembangan kegiatan dalam perekonomian yang menyebabkan barang dan jasa yang diproduksi dalam masyarakat bertambah dan kemakmuran masyarakat meningkat. Dalam APBN tahun 2016 ditetapkan bahwa target laju pertumbuhan ekonomi yang ingin dicapai

adalah sebesar 5.3 persen. Sehingga, pada penelitian ini kategori 1 merupakan kabupaten/kota yang memiliki laju pertumbuhan ekonomi lebih dari sama dengan 5.3 persen. Sedangkan kategori 0, merupakan kabupaten/kota yang memiliki laju pertumbuhan ekonomi kurang dari 5.3 persen.

15. Daerah ekonomi dibedakan menjadi dua, yaitu daerah ekonomi maju dan ekonomi kurang maju. Daerah ekonomi maju merupakan daerah yang masyarakat serta wilayahnya telah berkembang dalam skala nasional. Sedangkan daerah ekonomi kurang maju merupakan daerah yang masyarakat dan wilayahnya relatif kurang berkembang dibandingkan daerah lain dalam skala nasional. Daerah ekonomi maju dikelompokkan ke dalam zona 1, zona 2, dan zona 3. Sedangkan daerah ekonomi kurang maju dikelompokkan ke dalam zona 4 dan zona 5.
16. Sektor penyumbang PDRB terbesar, yaitu sektor yang memberikan kontribusi paling besar dalam pembentukan PDRB di daerah tersebut. Dalam hal ini, peneliti menggunakan 4 kategori berdasarkan struktur pembentuk PDRB Provinsi Jawa Timur tahun 2016 menurut lapangan usaha yang mendominasi.

3.3 Langkah Analisis

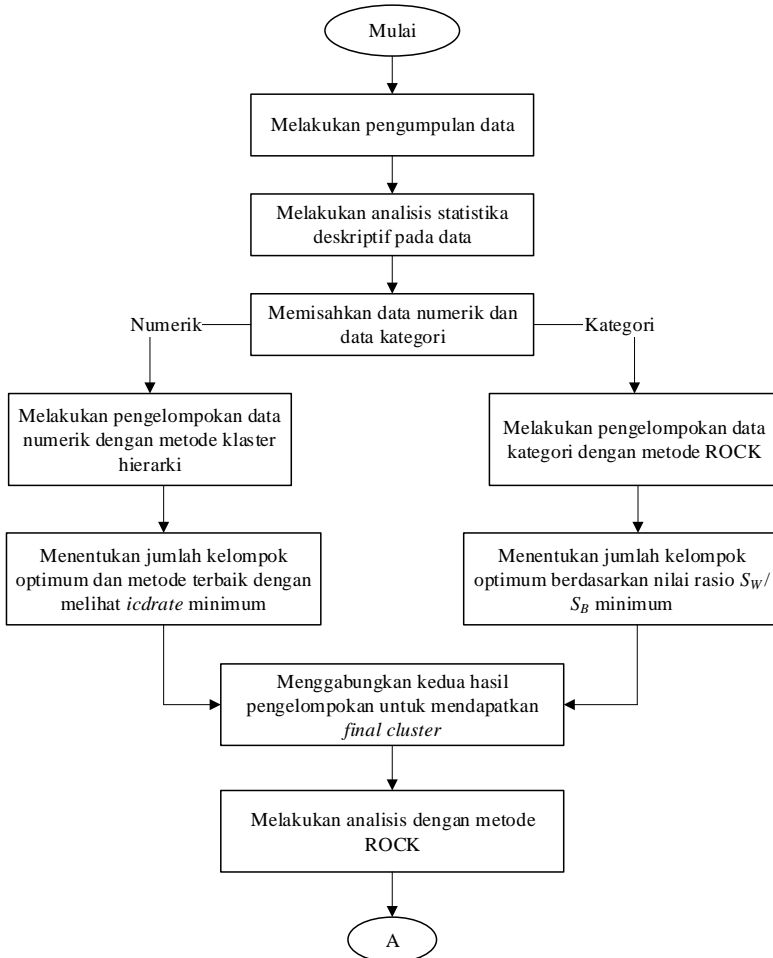
Berikut merupakan langkah analisis yang dilakukan dalam penelitian ini.

1. Melakukan pengumpulan data sesuai dengan indikator pembangunan kualitas SDM dan indikator pembangunan ekonomi pada masing-masing kabupaten/kota di Provinsi Jawa Timur tahun 2016.
2. Mendeskripsikan karakteristik dari masing-masing kabupaten/kota di Provinsi Jawa Timur berdasarkan data yang diperoleh.
3. Membagi data menjadi dua bagian, yaitu dari 21 variabel campuran numerik kategori menjadi 17 variabel berskala numerik dan 4 variabel berskala kategori.
4. Melakukan pengelompokan pada data berskala numerik menggunakan metode hierarki *agglomerative*. Jarak yang digunakan

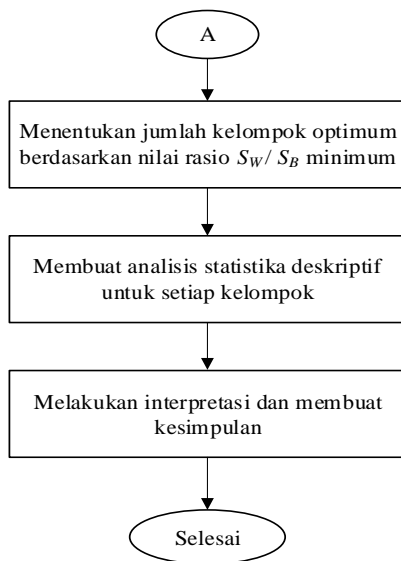
adalah jarak *Ecludiean* dan metode yang digunakan adalah *single linkage*, *complete linkage*, *average linkage*, dan *Ward's*.

5. Melakukan perhitungan nilai *Pseudo-F* dan *icdrate* untuk menentukan metode pengelompokan dengan jumlah kelompok paling optimum pada metode hierarki *agglomerative*.
6. Melakukan pengelompokan pada data berskala kategori menggunakan metode ROCK. Nilai θ yang digunakan dalam penelitian ini adalah 0.05; 0,07; 0.1; 0.12; 0.15; 0.17; 0.2; 0.22; 0.25; 0.27; dan 0.3.
7. Melakukan perhitungan rasio S_W dan S_B untuk menentukan jumlah kelompok yang optimum pada pengelompokan data kategori yang disesuaikan dengan objek pengamatan dan hasil pengelompokan yang diharapkan.
8. Melakukan penggabungan hasil pengelompokan data numerik dan kategori berdasarkan metode *ensemble* ROCK. Kemudian hasil penggabungan tersebut dikelompokkan kembali menggunakan metode ROCK untuk menghasilkan *final cluster*.
9. Melakukan perhitungan rasio S_W dan S_B untuk menentukan jumlah kelompok yang optimum pada pengelompokan data campuran numerik dan kategori.
10. Mendeskripsikan karakteristik dari masing-masing kelompok berdasarkan hasil pengelompokan data campuran numerik dan kategori.
11. Membuat kesimpulan dan saran.

Langkah analisis dalam penelitian ini dapat digambarkan secara rinci melalui diagram alir sebagai berikut.



Gambar 3.1 Diagram Alir Penelitian



Gambar 3.1 Diagram Alir Penelitian (lanjutan)

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab ini, dilakukan analisis dan pembahasan terkait karakteristik data dan pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator pembangunan SDM dan pembangunan ekonomi.

4.1 Karakteristik Kabupaten/ Kota di Jawa Timur Berdasarkan Pembangunan Kualitas SDM dan Pembangunan Ekonomi

Provinsi Jawa Timur terdiri dari 38 kabupaten/kota yang masing-masing memiliki karakteristik yang berbeda. Oleh karena itu, dilakukan analisis mengenai karakteristik kabupaten/kota yang dibedakan berdasarkan skala variabel, yaitu numerik dan kategori.

4.1.1 Karakteristik Berdasarkan Skala Data Numerik

Berikut ini merupakan karakteristik kabupaten/kota di Jawa Timur berdasarkan data numerik.

Tabel 4.1 Karakteristik Data Numerik

Variabel	Mean	Median	St. Dev	Min	Max
Jumlah rumah sakit	9.76	6	10.57	1	62
Jumlah puskesmas	25.26	25	12.84	3	62
Persentase obat & vaksin	85.63	86.5	8.65	58	100
Persentase RT ber-PHBS	49.62	50.30	14.75	19.4	75.1
Persentase jaminan kesehatan	50.34	47.91	20.65	27.22	134.67
AHH	71.048	71.7	2.06	65.89	73.87
Jumlah ibu meninggal	14.05	13	8.33	1	37
AKB	28.39	24.32	11.44	15.6	56.13

Tabel 4.1 Karakteristik Data Numerik (lanjutan)

Variabel	Mean	Median	St. Dev	Min	Max
Angka kesakitan	16.12	15.66	3.24	9.84	25.09
Rasio murid-guru SD	13.65	13.49	3.63	7.34	21.88
Rasio murid-guru SMP	11.18	11.53	2.27	6.09	16.08
Rasio murid-guru SMA	9.50	9.23	2.32	4.09	15.85
APK SD	108.14	107.92	3.83	101.67	115.74
APK SMP	90.80	90.82	7	70.96	101.57
APK SMA	89.10	86.16	22.91	43.07	139.87
Rata-rata lama sekolah	7.470	7.22	1.7	3.790	11.09
UMK	1729158	1459000	563463	1283000	3045000

Rumah sakit dan puskesmas merupakan salah satu fasilitas kesehatan yang disediakan pemerintah untuk memenuhi kebutuhan masyarakat akan pelayanan kesehatan. Berdasarkan Tabel 4.1 terlihat bahwa di Provinsi Jawa Timur rata-rata memiliki jumlah rumah sakit sekitar 10 rumah sakit di tiap daerahnya. Namun, ada juga daerah yang hanya memiliki satu rumah sakit, yaitu Kabupaten Sampang dan Kota Pasuruan. Di sisi lain, Kota Surabaya merupakan daerah yang memiliki rumah sakit dengan jumlah paling banyak, yaitu sejumlah 62 rumah sakit. Sedangkan, rata-rata jumlah puskesmas di Jawa Timur adalah 25 puskesmas.

Variabel ketersediaan obat dan vaksin merupakan salah satu indikator untuk mengukur kinerja dari program kefarmasian dan alat kesehatan yang bertujuan untuk meningkatkan akses dan pelayanan prima di bidang kesehatan. Rata-rata ketersediaan obat dan vaksin di Jawa Timur adalah sebesar 85.63%. Sementara target yang ditetapkan oleh Dinas Kesehatan Provinsi Jawa Timur pada tahun 2016 adalah sebesar 80%. Sehingga dapat dikatakan bahwa

ketersediaan obat dan vaksin di sebagian besar kabupaten/kota di Jawa Timur telah mencapai target. Namun dari Tabel 4.1 juga terlihat bahwa masih terdapat kabupaten/kota yang memiliki persentase ketersediaan obat dan vaksin sebesar 58%. Oleh karena itu, kegiatan yang dirasa perlu untuk meningkatkan persentase ketersediaan obat dan vaksin di kabupaten/kota yang berada jauh di bawah target perlu dilakukan agar terjadi pemerataan pelayanan kesehatan. Jika dilihat dari persentase RT ber-PHBS di Jawa Timur, rata-rata sebesar 49.62% dengan persentase terendah berada di Kabupaten Bondowoso, yaitu sebesar 19.4%. Sedangkan, rata-rata persentase masyarakat Jawa Timur yang memiliki jaminan kesehatan adalah 50.34%. Hal ini menunjukkan bahwa kesadaran masyarakat Jawa Timur akan pentingnya kesehatan serta perilaku hidup bersih dan sehat masih sangat kurang. Rata-rata indikator kesehatan lainnya seperti AHH, jumlah ibu meninggal, AKB, dan angka kesakitan dapat dilihat pada Tabel 4.1.

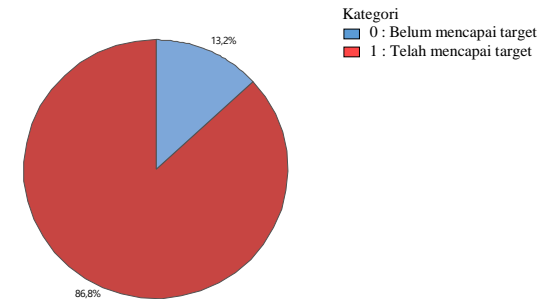
Kualitas SDM di Jawa Timur, jika dilihat dari segi pendidikan dapat tercermin melalui variabel rasio murid-guru dan APK untuk jenjang pendidikan SD, SMP, SMA serta rata-rata lama sekolah. Untuk menciptakan kondisi belajar yang ideal, diperlukan perbandingan antara murid dan guru yang sesuai. Semakin tinggi nilai rasio murid-guru, berarti semakin berkurang tingkat pengawasan dan perhatian guru terhadap murid. Sehingga mutu pengajaran juga cenderung semakin rendah. Berdasarkan Tabel 4.1 terlihat bahwa rata-rata rasio murid-guru untuk jenjang pendidikan SD, SMP, SMA masing-masing adalah sebesar 13.65, 11.18, dan 9.5. Nilai tersebut menunjukkan bahwa di Provinsi Jawa Timur semakin tinggi jenjang pendidikannya, maka mutu pengajaran yang diberikan cenderung semakin tinggi. Besarnya APK menunjukkan tingkat partisipasi penduduk secara umum pada suatu tingkat pendidikan tanpa memperhatikan ketepatan usia sekolah pada jenjang pendidikannya. Nilai APK yang mendekati atau melebihi 100 persen menunjukkan bahwa ada penduduk yang sekolah belum mencukupi umur dan atau melebihi umur seharusnya. Rata-rata APK untuk jenjang pendidikan SD, SMP, dan SMA di Provinsi

Jawa Timur masing-masing adalah sebesar 108.14, 90.80, dan 89.10. Rata-rata APK pada jenjang pendidikan SD yang melebihi 100 tersebut menunjukkan bahwa rata-rata penduduk Jawa Timur yang bersekolah pada jenjang pendidikan SD tidak sesuai dengan umur seharusnya. Jika dilihat dari variabel rata-rata lama sekolah, pada tahun 2016 Provinsi Jawa Timur memiliki rata-rata lama sekolah sebesar 7.47 atau sekitar 8 tahun. Selain itu, rata-rata lama sekolah tertinggi di Jawa Timur hanya sebesar 11.08 atau 11 tahun. Hal ini menunjukkan bahwa program pendidikan 12 tahun yang digagas oleh pemerintah masih belum terpenuhi untuk Provinsi Jawa Timur pada tahun 2016.

Penetapan upah minimum oleh Gubernur dilakukan berdasarkan Kebutuhan Hidup Layak (KHL) dengan memperhatikan produktivitas dan pertumbuhan ekonomi daerah. Artinya semakin tinggi UMK suatu daerah maka semakin tinggi pula tingkat produktivitas dan pertumbuhan ekonomi daerah tersebut. Pada Tabel 4.1 terlihat bahwa rata-rata UMK Provinsi Jawa Timur adalah sebesar Rp.1.729.158. Namun, jika dilihat nilai UMK terendah dan tertinggi di Provinsi Jawa Timur memiliki *range* yang cukup besar. Hal ini menunjukkan bahwa pemerataan ekonomi di Provinsi Jawa Timur masih cukup rendah. Nilai

4.1.2 Karakteristik Berdasarkan Skala Data Kategori

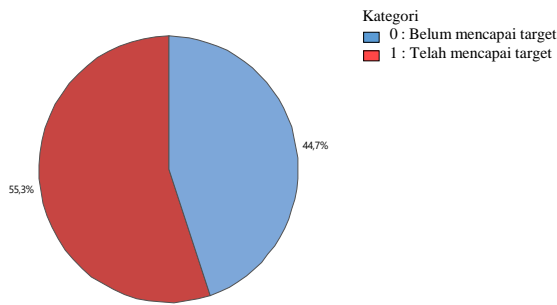
Variabel kategori pada penelitian ini terdiri dari variabel indeks *gini ratio*, laju pertumbuhan ekonomi, daerah ekonomi, dan sektor penyumbang PDRB terbesar. Dalam APBN 2016 tercantum mengenai beberapa indikator dan target pembangunan nasional, antara lain pertumbuhan ekonomi sebesar 5.3% dan indeks *gini ratio* sebesar 0.39. Oleh karena itu dalam penelitian ini dilakukan pengkategorian terhadap kabupaten/kota sesuai dengan pencapaian pada masing-masing indikator. Bagi kabupaten/kota yang telah mencapai target dikategorikan ke dalam kategori satu (1) dan yang belum mencapai target dikategorikan dalam kategori nol (0), sesuai yang tercantum dalam APBN 2016. Berikut ini merupakan karakteristik Provinsi Jawa Timur berdasarkan keempat variabel kategori tersebut.



Gambar 4.1 Diagram Lingkaran Indeks *Gini Ratio*

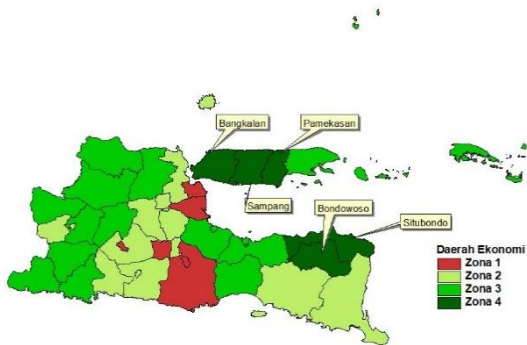
Berdasarkan Gambar 4.1 terlihat bahwa sebesar 86.8% kabupaten/kota di Jawa Timur telah mencapai target. Artinya lebih dari 50 persen daerah di Provinsi Jawa Timur memiliki indeks *gini ratio* kurang dari 0.39. Hal ini menunjukkan bahwa sebagian besar kabupaten/kota di Jawa Timur tergolong dalam ketimpangan pendapatan yang rendah. Sedangkan 13.2% kabupaten/kota lainnya tergolong dalam ketimpangan pendapatan sedang. Nilai *gini ratio* yang semakin tinggi atau mendekati satu, menunjukkan bahwa ketimpangan pendapatan yang terjadi di daerah tersebut juga semakin tinggi.

Sementara itu, apabila dilihat dari variabel laju pertumbuhan ekonomi. Sebanyak 55.3% kabupaten/kota di Jawa Timur telah mencapai target pembangunan. Namun, 44.7% sisanya belum mencapai target pembangunan. Hal ini menunjukkan bahwa sebanyak 55.3% kabupaten/kota di Jawa Timur mengalami pertumbuhan ekonomi di atas 5.3 persen dan 44.7% sisanya masih di bawah 5.3 persen. Laju pertumbuhan ekonomi dapat diukur dari laju pertumbuhan PDRB suatu daerah. Semakin tinggi laju pertumbuhan ekonomi suatu daerah, artinya semakin tinggi pula tingkat keberhasilan pembangunan di daerah tersebut dalam periode waktu tertentu. Pertumbuhan ekonomi yang positif menunjukkan adanya kenaikan produksi barang dan jasa. Berikut ini merupakan diagram lingkaran yang menggambarkan kondisi laju pertumbuhan ekonomi di kabupaten/kota di Jawa Timur pada tahun 2016.



Gambar 4.2 Diagram Lingkaran Laju Pertumbuhan Ekonomi

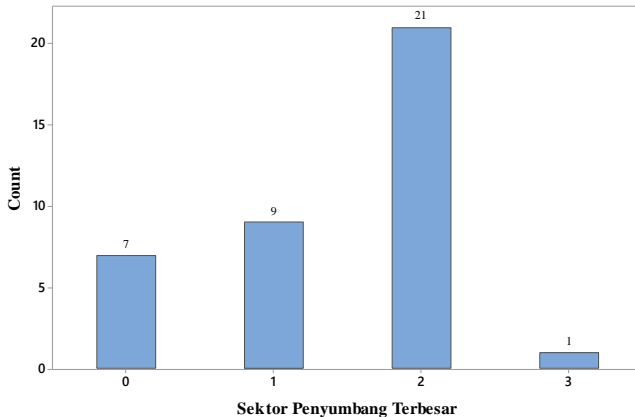
Pembagian daerah ekonomi maju dan daerah ekonomi kurang maju tercantum dalam Peraturan Menteri Komunikasi dan Informatika Republik Indonesia. Daerah ekonomi maju dan daerah ekonomi kurang maju ditetapkan berdasarkan Indeks Potensi dan Kemajuan Daerah (IPKD) yang terbagi dalam 5 zona. Zona 1, zona 2, dan zona 3 merupakan daerah-daerah yang termasuk dalam daerah ekonomi maju. Sedangkan daerah kurang maju berada pada zona 4 dan zona 5. Berikut ini merupakan peta Provinsi Jawa Timur berdasarkan pembagian daerah ekonomi.



Gambar 4.3 Peta Daerah Ekonomi

Berdasarkan Gambar 4.3 terlihat bahwa kabupaten/kota di Provinsi Jawa Timur tidak ada yang berada pada zona 5. Namun, terdapat 5 daerah yang berada pada zona 4 atau yang disebut sebagai daerah ekonomi kurang maju. 5 daerah yang berada di zona

4 adalah Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Bangkalan, Kabupaten Bondowoso, dan Kabupaten Situbondo. Sebanyak 33 kabupaten/kota lainnya berada di zona 1, 2, atau 3. Sehingga dapat dikatakan bahwa sebagian besar kabupaten/kota di Jawa Timur telah tergolong sebagai daerah ekonomi maju.



Gambar 4.4 Diagram Batang Sektor Penyumbang PDRB Terbesar

Pada Gambar 4.4 dapat dilihat bahwa sektor penyumbang PDRB terbesar pada sebagian besar kabupaten/kota di Provinsi Jawa Timur pada tahun 2016 adalah sektor pertanian, kehutanan, dan perikanan. Kemudian diikuti oleh sektor perdagangan besar dan eceran, reparasi mobil dan sepeda motor dengan 9 kabupaten/kota dan sektor industri pengolahan dengan 7 kabupaten/kota. Sementara 1 kabupaten yang disumbang oleh sektor lainnya adalah Kabupaten Bojonegoro.

4.2 Pengelompokan Kabupaten/ Kota di Jawa Timur Berdasarkan Pembangunan Kualitas SDM dan Pembangunan Ekonomi

Tahapan pengelompokan data campuran, dimulai dengan mengelompokkan data secara terpisah sesuai dengan jenis skala data. Berikut ini merupakan hasil pengelompokan data numerik, pengelompokan data kategori, dan pengelompokan data campuran.

4.2.1 Pengelompokan Data Numerik

Data numerik yang digunakan dalam penelitian ini dapat dilihat pada Tabel 3.1, yaitu variabel $X_1 - X_{17}$. Sedangkan, metode pengelompokan yang digunakan untuk mengelompokkan data numerik adalah metode hierarki *agglomerative* dengan ukuran jarak *euclidean*. Metode hierarki *agglomerative* yang digunakan meliputi, metode *single linkage*, *complete linkage*, *average linkage*, dan *ward's method*. Pengelompokan dilakukan dengan membentuk 2 hingga 4 kluster. Kemudian dari masing-masing kluster yang terbentuk di setiap metode dibandingkan untuk mendapatkan hasil pengelompokan yang optimum. Pemilihan jumlah kelompok optimum dilakukan dengan melihat nilai *Pseudo-F* terbesar pada masing-masing metode kluster. Kemudian nilai *icdrate* (*internal cluster dispersion rate*) dari jumlah kelompok optimum pada masing-masing metode dibandingkan untuk mendapatkan hasil pengelompokan terbaik. Hasil pengelompokan terbaik dipilih berdasarkan jumlah kelompok yang menghasilkan nilai *icdrate* terkecil. Berikut ini merupakan nilai *Pseudo-F* dan *icdrate* yang didapatkan dari perhitungan dengan Persamaan 2.16 dan Persamaan 2.17.

Tabel 4.2 Hasil Kinerja Pengelompokan Data Numerik

Metode	Jumlah Kelompok	<i>Pseudo-F</i>	<i>Icdrate</i>
<i>Single linkage</i>	2	4.5919	0.88694
	3	4.0259	0.81312
	4	3.7136	0.75336
<i>Complete linkage</i>	2	7.6167	0.8255
	3	9.363	0.65164
	4	8.7619	0.56421
<i>Average linkage</i>	2	4.5919	0.88694
	3	4.9575	0.77933
	4	7.3287	0.6074
<i>Ward's method</i>	2	8.8952	0.80194
	3	9.363	0.65164
	4	9.4766	0.54482

Berdasarkan Tabel 4.2 dapat dilihat bahwa berdasarkan kriteria nilai *Pseudo-F* terbesar dan *icdrate* terkecil dihasilkan oleh

pengelompokan menggunakan *ward's method* dengan jumlah kelompok sebanyak 4. Berdasarkan hasil pengelompokan tersebut, masing-masing kelompok terdiri dari 20 kabupaten/kota, 9 kabupaten/kota, 3 kabupaten/kota, dan 6 kabupaten/kota. Daftar anggota dari masing-masing kelompok dapat dilihat pada Lampiran 6.

4.2.2 Pengelompokan Data Kategori

Data kategori yang digunakan dalam pengelompokan ini adalah variabel $X_{18} - X_{21}$, sesuai yang tercantum dalam Tabel 3.2. Sedangkan metode pengelompokan yang digunakan untuk data berskala kategori adalah metode ROCK dengan *threshlod* yang digunakan sebesar 0.05, 0.1, 0.12, 0.15, 0.17, 0.2, 0.22, 0.25, 0.27, dan 0.3. Hasil yang diharapkan dalam pengelompokan ini adalah tidak ada kelompok yang berisi satu anggota dan semua objek pengamatan tidak berada dalam satu kelompok. Pemilihan kelompok optimum pada metode ROCK dilakukan dengan membandingkan nilai rasio S_w/S_B yang dihasilkan. Kelompok yang optimum adalah kelompok dengan kriteria nilai rasio S_w/S_B terkecil. Berikut merupakan hasil pengelompokan data kategori berdasarkan nilai rasio S_w/S_B yang didapatkan dari perhitungan Persamaan 2.24 dibagi dengan perhitungan Persamaan 2.25.

Tabel 4.3 Hasil Kinerja Pengelompokan Data Kategori

<i>Threshold</i>	Jumlah Kelompok	Rasio S_w/S_B
0.05	2	0.2361
0.07	2	0.2361
0.1	2	0.0778
0.12	2	0.0943
0.15	2	0.1012
	3	0.1082
	4	0.1290
0.17	2	0.1012
	3	0.1082
	4	0.1533
0.2	2	0.0455
	3	0.1321

Tabel 4.3 Hasil Kinerja Pengelompokan Data Kategori (lanjutan)

<i>Threshold</i>	Jumlah Kelompok	Rasio S_w/S_B
0.2	4	0.1253
0.22	2	2.87×10^{-16}
	3	0.1321
	4	0.1608
0.25	2	0.0710
	3	0.1554
0.27	2	0.0710
	3	0.1562
0.3	2	0.0710
	3	0.1341

Berdasarkan Tabel 4.3 tersebut terlihat bahwa kelompok optimum dihasilkan oleh *threshold* 0.22 dengan jumlah kelompok yang terbentuk adalah 2 kelompok dan nilai rasio S_w/S_B sebesar 2.87×10^{-16} . Kelompok pertama terdiri dari 17 kabupaten/kota dan kelompok kedua terdiri dari 21 kabupaten/kota. Anggota dari masing-masing kelompok dapat dilihat pada Lampiran 7. Sedangkan nilai rasio S_w/S_B pada pengelompokan data kategori dengan nilai *threshold* yang berbeda dapat dilihat pada Lampiran 9

4.2.3 Pengelompokan Data Campuran

Setelah mendapatkan hasil pengelompokan dari masing-masing pengelompokan berdasarkan skala data. Kemudian hasil pengelompokan tersebut digabungkan dan dikelompokkan kembali menggunakan metode ROCK. Berikut ini merupakan rangkuman dari hasil pengelompokan data numerik dan pengelompokan data kategori.

Tabel 4.4 Rangkuman Hasil Pengelompokan Data Numerik dan Kategori

No.	Kabupaten/ kota	Pengelompokan Data Numerik	Pengelompokan Data Kategori
1	Kabupaten Pacitan	1	1
2	Kabupaten Ponorogo	1	1
3	Kabupaten Trenggalek	1	1
4	Kabupaten Tulungagung	1	1
5	Kabupaten Blitar	1	1

Tabel 4.4 Rangkuman Hasil Pengelompokan Data Numerik dan Kategori
(lanjutan)

No.	Kabupaten/ kota	Pengelompokan Data Numerik	Pengelompokan Data Kategori
6	Kabupaten Kediri	1	1
7	Kabupaten Malang	1	2
8	Kabupaten Lumajang	1	1
9	Kabupaten Jember	2	1
10	Kabupaten Banyuwangi	1	2
11	Kabupaten Bondowoso	2	1
12	Kabupaten Situbondo	2	1
13	Kabupaten Probolinggo	2	1
14	Kabupaten Pasuruan	2	2
15	Kabupaten Sidoarjo	3	2
16	Kabupaten Mojokerto	1	2
17	Kabupaten Jombang	1	2
18	Kabupaten Nganjuk	1	1
19	Kabupaten Madiun	1	1
20	Kabupaten Magetan	1	2
21	Kabupaten Ngawi	1	1
22	Kabupaten Bojonegoro	1	2
23	Kabupaten Tuban	1	1
24	Kabupaten Lamongan	1	2
25	Kabupaten Gresik	3	2
26	Kabupaten Bangkalan	2	1
27	Kabupaten Sampang	2	2
28	Kabupaten Pamekasan	2	2
29	Kabupaten Sumenep	2	1
30	Kota Kediri	4	2
31	Kota Blitar	4	2
32	Kota Malang	1	2
33	Kota Probolinggo	1	2
34	Kota Pasuruan	1	2
35	Kota Mojokerto	4	2
36	Kota Madiun	4	2
37	Kota Surabaya	3	2
38	Kota Batu	4	2

Hasil pengelompokan dari masing-masing variabel dengan skala data yang berbeda pada Tabel 4.4, kemudian dinyatakan

sebagai variabel baru dengan skala data kategori. Variabel A merupakan variabel yang berisikan hasil pengelompokan dari data numerik. Sedangkan variabel B merupakan variabel yang berisikan hasil pengelompokan data kategori. Kedua variabel tersebut kemudian dikelompokkan kembali dengan metode ROCK sesuai dengan algoritma CEBMDC yang tercantum pada Gambar 2.1.

Berikut ini merupakan struktur data baru yang digunakan dalam pengelompokan data campuran.

Tabel 4.5 Struktur Data Baru

No.	Kabupaten/ kota	A (Pengelompokan Data Numerik)	B (Pengelompokan Data Kategori)
1	Kabupaten Pacitan	1	1
2	Kabupaten Ponorogo	1	1
3	Kabupaten Trenggalek	1	1
⋮	⋮	⋮	⋮
38	Kota Batu	4	2

Dengan struktur data yang baru sesuai dengan Tabel 4.5 di atas kemudian dilakukan pengelompokan untuk mendapatkan *final cluster*. Nilai *threshold* yang digunakan dalam pengelompokan tersebut adalah 0.05, 0.1, 0.12, 0.15, 0.17, 0.2, 0.22, 0.25, 0.27, dan 0.3. Rasio S_w/S_B adalah indeks yang digunakan sebagai acuan dalam menentukan jumlah kelompok optimum. Kelompok yang paling optimum adalah kelompok dengan nilai rasio S_w/S_B terkecil. Hasil perhitungan rasio S_w/S_B sesuai pada persamaan 2.24 dan 2.25 pada pengelompokan data campuran yang telah dilakukan dapat dilihat pada Tabel 4.6. Pada Tabel 4.6 terlihat bahwa pengelompokan dengan nilai *threshold* 0.05, 0.1, 0.12, 0.15, 0.17, 0.2, 0.22, 0.25, 0.27, dan 0.3 dengan jumlah kelompok yang terbentuk adalah 2 kelompok menghasilkan nilai rasio S_w/S_B yang cenderung konstan, yaitu sebesar 2.98×10^{-16} . Kemudian, dilakukan pengelompokan kembali menggunakan nilai *threshold* yang sama. Namun, pada pengelompokan kali ini jumlah kelompok yang diharapkan untuk terbentuk adalah 3 kelompok. Nilai rasio S_w/S_B yang dihasilkan

pada pengelompokan tersebut juga cenderung konstan, yaitu sebesar 0.1272. Nilai rasio S_w/S_B pada pengelompokan dengan nilai *threshold* yang berbeda dapat dilihat pada Lampiran 10.

Tabel 4.6 Hasil Kinerja Pengelompokan Data Campuran

<i>Threshold</i>	Jumlah Kelompok	Rasio S_w/S_B
0.05	2	2.98×10^{-16}
	3	0.1272
0.1	2	2.98×10^{-16}
	3	0.1272
0.12	2	2.98×10^{-16}
	3	0.1272
0.15	2	2.98×10^{-16}
	3	0.1272
0.17	2	2.98×10^{-16}
	3	0.1272
0.2	2	2.98×10^{-16}
	3	0.1272
0.22	2	2.98×10^{-16}
	3	0.1272
0.25	2	2.98×10^{-16}
	3	0.1272
0.27	2	2.98×10^{-16}
	3	0.1272
0.3	2	2.98×10^{-16}
	3	0.1272

Berdasarkan pengelompokan yang telah dilakukan, baik dengan 2 kelompok maupun 3 kelompok menunjukkan bahwa pengelompokan dengan membentuk 2 kelompok adalah kelompok yang optimum. Hal ini dibuktikan dari nilai rasio S_w/S_B pada 2 kelompok yang lebih kecil dari pada 3 kelompok. Selain itu, anggota dari 2 kelompok yang terbentuk baik dengan *threshold* 0.05, 0.1, hingga 0.3 adalah sama. Sehingga pemilihan *threshold* 0.05 atau 0.1 hingga 0.3 tidak memberikan efek yang berbeda. *Final cluster* yang didapatkan dari pengelompokan data campuran adalah 2 kelompok dengan masing-masing kelompok beranggota-

kan 17 kabupaten/kota dan 21 kabupaten/kota. Daftar anggota dari masing-masing kelompok dapat dilihat pada Tabel 4.7 berikut.

Tabel 4.7 Daftar Anggota Hasil Pengelompokan Data Campuran

Kelompok	Anggota Kelompok
1	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Lumajang, Jember, Bondowoso, Situbondo, Probolinggo, Nganjuk, Madiun, Ngawi, Tuban, Bangkalan, dan Sumenep.
2	Malang, Banyuwangi, Pasuruan, Sidoarjo, Mojokerto, Jombang, Magetan, Bojonegoro, Lamongan, Gresik, Sampang, Pamekasan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu.

4.2.4 Karakteristik Setiap Kelompok

Hasil pengelompokan pada data campuran yang dilakukan menggunakan metode *ensemble* ROCK dengan nilai *threshold* sebesar 0.05 dan menghasilkan 2 kelompok optimum dengan karakteristik setiap kelompok sebagai berikut.

a. Karakteristik Kelompok Berdasarkan Data Numerik

Karakteristik untuk masing-masing kelompok berdasarkan data numerik ditunjukkan oleh tabel berikut.

Tabel 4.8 Karakteristik Kelompok Satu Berdasarkan Data Numerik

Variabel	Mean	Median	St. Dev	Min	Max
X ₁	6.059	5	3.17	3	12
X ₂	27.82	25	7.58	17	49
X ₃	88.47	88	7.13	72	100
X ₄	47.01	45.80	16.41	19.40	71.20
X ₅	41.44	40.19	9.4	27.22	60.38
X ₆	70.44	70.67	2.2	65.89	73.4
X ₇	14.71	12	6.69	6	33
X ₈	32.25	27.31	13.23	18.01	56.13
X ₉	15.72	15.01	3.27	11.87	24.53
X ₁₀	11.62	11.50	2.73	7.34	17.19
X ₁₁	10.97	11.01	2.07	7.22	15.61
X ₁₂	9.75	9.22	2.49	4.09	14.45

Tabel 4.8 Karakteristik Kelompok Satu Berdasarkan Data Numerik (lanjutan)

Variabel	Mean	Median	St. Dev	Min	Max
X ₁₃	109.45	109.29	3.54	101.67	115.74
X ₁₄	89.31	90.15	7.82	70.96	100.50
X ₁₅	77.30	81.61	15.86	43.07	101.17
X ₁₆	6.47	6.54	0.85	5.08	7.73
X ₁₇	1433941	1411000	143274	1283000	1757000

Tabel 4.8 tersebut menjelaskan tentang karakteristik dari kelompok satu berdasarkan setiap variabelnya. Karakteristik dari kelompok satu maupun kelompok dua digunakan untuk menggambarkan dampak dari pembangunan SDM dan pembangunan ekonomi yang telah dilakukan pemerintah selama ini. Selanjutnya untuk karakteristik dari kelompok dua berdasarkan data numerik dapat dijelaskan oleh Tabel 4.9 berikut.

Tabel 4.9 Karakteristik Kelompok Dua Berdasarkan Data Numerik

Variabel	Mean	Median	St. Dev	Min	Max
X ₁	12.76	10	13.32	1	62
X ₂	23.19	22	15.78	3	62
X ₃	83.33	85	9.24	58	95
X ₄	51.74	52.7	13.29	28.7	75.10
X ₅	57.55	50.69	24.42	29.45	134.67
X ₆	71.54	72.05	1.84	66.95	73.87
X ₇	13.52	14	9.59	1	37
X ₈	25.28	20.26	8.9	15.60	43.65
X ₉	16.44	16.17	3.26	9.84	25.09
X ₁₀	15.29	15.14	3.47	9.68	21.88
X ₁₁	11.36	11.62	2.46	6.09	16.08
X ₁₂	9.31	9.24	2.2	6.25	15.85
X ₁₃	107.08	106.44	3.81	101.90	115.16
X ₁₄	92	90.95	6.18	74.63	101.57
X ₁₅	98.64	99.64	23.57	54.05	139.87
X ₁₆	8.28	8.45	1.8	3.79	11.09
X ₁₇	1968143	1603000	660974	1283000	3045000

Berdasarkan Tabel 4.8 dan Tabel 4.9 tersebut perbandingan nilai *mean* dari karakteristik masing-masing kelompok berdasarkan data numerik dapat dilihat pada tabel 4.10.

Tabel 4.10 Perbandingan Nilai *Mean* Setiap Kelompok

Variabel	Kelompok Satu	Kelompok Dua
X ₁	6.059	12.76
X ₂	27.82	23.19
X ₃	88.47	83.33
X ₄	47.01	51.74
X ₅	41.44	57.55
X ₆	70.44	71.54
X ₇	14.71	13.52
X ₈	32.25	25.28
X ₉	15.72	16.44
X ₁₀	11.62	15.29
X ₁₁	10.97	11.36
X ₁₂	9.75	9.31
X ₁₃	109.45	107.08
X ₁₄	89.31	92
X ₁₅	77.30	98.64
X ₁₆	6.47	8.28
X ₁₇	1433941	1968143

*Catatan : huruf tebal menandakan nilai yang lebih baik

Pada Tabel 4.10 dapat dilihat bahwa berdasarkan nilai *mean* data numerik pada kelompok dua, cenderung memiliki kriteria yang lebih baik dari pada karakteristik kelompok satu.

b. Karakteristik Kelompok Berdasarkan Data Kategori

Karakteristik dari masing-masing kelompok berdasarkan data kategori yang diwakili oleh modus dari variabel X₁₈ – X₂₁ dapat dilihat pada gambar berikut.

Tabel 4.11 Karakteristik Setiap Kelompok Berdasarkan Data Kategori

Variabel	Kelompok Satu	Kelompok Dua
X ₁₈	1	1
X ₁₉	0	1
X ₂₀	3	2
X ₂₁	3	2

Tabel 4.11 merupakan tabel yang menunjukkan karakteristik dari masing-masing kelompok berdasarkan kategori yang mendominasi dari masing-masing kelompok yang terbentuk.

Berdasarkan Tabel 4.10 dan Tabel 4.11 tersebut dapat disimpulkan bahwa karakteristik dari masing-masing kelompok berdasarkan data numerik dan data kategori adalah sebagai berikut.

1. Kabupaten/kota kurang maju

Kelompok satu berisikan 17 kabupaten/kota yang memiliki kondisi dari segi pembangunan SDM dan pembangunan ekonomi yang kurang baik. Jika dilihat dari fasilitas kesehatan yang dimiliki kelompok 1, rata-rata jumlah rumah sakit yang dimiliki cenderung lebih sedikit dari pada kelompok dua, yaitu 13 rumah sakit di setiap kabupaten/kota. Disisi lain, dapat dikatakan pula bahwa kesadaran masyarakat pada kelompok ini akan pentingnya kesehatan masih sangat kurang. Hal ini terbukti dari persentase RT yang berperilaku hidup bersih dan sehat serta persentase kepemilikan jaminan kesehatan yang jauh lebih rendah dari pada kelompok dua, yaitu sekitar 47.01 persen dan 41.44 persen. Meskipun rata-rata persentase ketersediaan obat dan vaksin di kelompok satu cenderung lebih banyak dari pada kelompok dua. Namun, kurangnya kesadaran masyarakat akan pentingnya kesehatan dapat menjadi salah satu faktor yang menyebabkan tingginya rata-rata jumlah ibu yang meninggal dan AKB di kelompok satu. Rata-rata AHH yang dimiliki oleh kelompok satu adalah 70.44 dengan rata-rata angka kesakitan sekitar 15.72. Dari segi pendidikan, rasio murid guru jenjang pendidikan SD dan SMA di kelompok satu cenderung lebih banyak dari pada kelompok dua. Tingginya rasio murid guru menggambarkan kurangnya mutu pengajaran di daerah tersebut. Selain itu, rata-rata lama sekolah yang dimiliki oleh kelompok satu adalah 6.47 atau sekitar 6 atau 7 tahun. Artinya tingkat pendidikan yang dimiliki masyarakat di kabupaten/kota yang tergolong kelompok satu masih cenderung rendah. Apabila dilihat dari variabel indeks *gini ratio*, terlihat bahwa sebagian besar anggota dari kelompok satu telah mencapai target pembangunan. Namun, apabila dilihat dari karakteristik laju pertumbuhan ekonomi, kelompok satu menunjukkan bahwa sebagian besar anggota dari kelompok satu masih belum

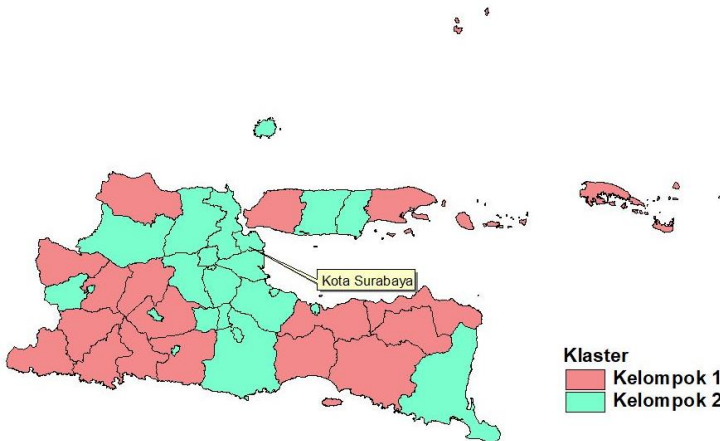
mencapai target pembangunan. Adanya pertumbuhan ekonomi apabila tidak disertai dengan peningkatan kesejahteraan bagi masyarakat miskin dapat berpotensi menimbulkan kerawanan sosial, begitu pula sebaliknya. Oleh karena itu, meskipun *gini ratio* pada kelompok satu telah mencapai target tetap perlu dilakukan usaha guna meningkatkan laju pertumbuhan ekonomi daerahnya. Sehingga masyarakat di setiap kabupaten/kota dapat merasakan standar hidup yang lebih baik. Berdasarkan Tabel 4.6 terlihat bahwa kelompok satu didominasi oleh kabupaten/kota yang memiliki kontribusi PDRB cukup tinggi di sektor pertanian, kehutanan, dan perikanan. Sektor pertanian, kehutanan, dan perikanan memiliki peranan yang besar dalam memajukan perekonomian bangsa. Sebagai salah satu sektor pokok, seharusnya kontribusi yang diberikan sektor pertanian, kehutanan, dan perikanan mampu menjadikan perekonomian daerah yang bersangkutan menjadi lebih baik. Namun, fakta menunjukkan bahwa perekonomian di kelompok satu masih relatif di bawah kelompok dua. Artinya, produktivitas di sektor tersebut masih belum mampu mendorong perekonomian daerah. Laju pertumbuhan ekonomi yang rendah menjadi salah satu penyebab kurangnya standar hidup layak pada masyarakat di daerah yang tergolong dalam kelompok satu. Sehingga berdampak pula pada rendahnya UMK di kabupaten/kota pada kelompok tersebut apabila dibandingkan dengan kelompok dua.

2. Kabupaten/kota maju

Kelompok kedua merupakan kelompok dengan kabupaten/kota yang memiliki kualitas yang lebih baik apabila dibandingkan dengan kelompok satu. Baik dari segi kesehatan, pendidikan, maupun tingkat perekonomian. Namun, tidak bisa dipungkiri bahwa ada beberapa hal yang masih perlu mendapat perhatian khusus. Berdasarkan Tabel 4.7 dari segi kesehatan, kelompok dua memiliki AKB dan jumlah kematian yang relatif lebih rendah. Rata-rata AHH pada kelompok tersebut juga lebih tinggi dari pada kelompok satu. Dari segi pendidikan, rasio murid guru SMP di kelompok dua memiliki rata-rata lebih

tinggi dari pada kelompok satu. Artinya mutu pengajaran untuk jenjang pendidikan SMP masih kurang baik, karena jumlah murid tidak seimbang dengan jumlah guru yang tersedia. Selain itu, rata-rata lama sekolah di daerah-daerah yang ada pada kelompok dua juga belum mencapai 12 tahun, yaitu hanya 8.28 atau sekitar 8 sampai 9 tahun. Dapat dikatakan bahwa rata-rata lama sekolah di kelompok dua masih lebih baik apabila dibandingkan dengan kelompok satu. Namun, tetap saja hal ini masih belum mencapai tujuan pemerintah yang mencanangkan wajib belajar 12 tahun. Dari sisi perekonomian, kelompok dua didominasi oleh kabupaten atau kota yang tergolong daerah ekonomi maju dengan sektor perdagangan besar dan eceran, reparasi mobil dan sepeda motor sebagai sektor penyumbang PDRB terbesar.

Persebaran daerah dari masing-masing kelompok dapat dilihat pada Gambar 4.5.



Gambar 4.5 Peta Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Hasil Pengelompokan *Ensemble ROCK*

Pada Gambar 4.5 tersebut terlihat bahwa sebagian besar kabupaten/kota yang berada di sekitar Kota Surabaya merupakan anggota dari kelompok dua atau kelompok dengan kabupaten/kota

maju dari sisi pembangunan SDM dan ekonomi. Dapat dikatakan pula bahwa semakin jauh lokasi daerah tersebut dari pusat Ibu Kota Provinsi (Kota Surabaya), maka semakin rendah pula dampak dari pembangunan SDM dan ekonomi yang dirasakan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut.

1. Kondisi kabupaten/kota di Jawa Timur berdasarkan dampak pembangunan kualitas SDM dan pembangunan ekonomi pada tahun 2016, yaitu.
 - a. Berdasarkan pembangunan kualitas SDM, yang digambarkan oleh variabel X_1 - X_{17} menunjukkan bahwa masih banyak terjadi kesenjangan di setiap kabupaten/kota di Provinsi Jawa Timur. Hal ini ditunjukkan oleh nilai minimum dan maksimum dari statistika deskriptif untuk masing-masing variabel yang terpaut cukup jauh. Oleh karena itu dapat dikatakan bahwa pemerataan kesejahteraan di Provinsi Jawa Timur masih cukup rendah apabila ditinjau dari segi pembangunan kualitas SDM.
 - b. Pada penelitian ini pembangunan ekonomi digambarkan oleh variabel X_{18} - X_{21} . Hampir sebagian besar kabupaten/kota di Jawa Timur telah mencapai target nasional berdasarkan variabel indeks *gini ratio* dan laju pertumbuhan ekonomi. Sedangkan berdasarkan pembagian daerah ekonomi, kabupaten/kota di Jawa Timur tergolong pada zona 1 hingga 4 dari 5 zona. Sektor penyumbang PDRB terbesar pada sebagian besar kabupaten/kota di Provinsi Jawa Timur pada tahun 2016 adalah sektor pertanian, kehutanan, dan perikanan.
2. Hasil pengelompokan kabupaten/kota di Jawa Timur berdasarkan metode *ensemble* ROCK.
 - a. Pengelompokan berdasarkan data numerik (X_1 - X_{17}) menghasilkan kelompok optimum pada pengelompokan dengan *ward's method*. Nilai *icdrate* yang dihasilkan adalah sebesar 0.54482 dengan jumlah kelompok yang terbentuk adalah 4 kelompok.

- b. Pengelompokan berdasarkan data kategori (X_{18} - X_{21}) menghasilkan kelompok optimum pada nilai *threshold* sebesar 0.22 dan rasio S_W/S_B sebesar 2.87×10^{-16} . Jumlah kelompok optimum yang terbentuk adalah sebanyak 2 kelompok.
- c. *Final cluster* yang didapatkan dari metode *ensemble* ROCK, yaitu menggabungkan pengelompokan data numerik dengan pengelompokan data kategori adalah terbentuk kelompok optimum dengan jumlah kelompok sebanyak 2. Sedangkan penggunaan nilai *threshold* antara 0.05, 0.1, hingga 0.3 yang telah ditentukan sebelumnya tidak memberikan perbedaan yang berarti karena nilai rasio S_W/S_B yang dihasilkan cenderung konstan, yaitu 2.98×10^{-16} dengan anggota yang terbentuk juga sama.

Berdasarkan *final cluster*, hasil pengelompokan data campuran terbentuk sebanyak 2 kelompok. Kelompok 1 yang beranggotakan 17 kabupaten/kota diidentifikasi sebagai kelompok yang memiliki kondisi dari segi pembangunan kualitas SDM dan pembangunan ekonomi yang cenderung lebih rendah dari pada kelompok 2 yang beranggotakan 21 kabupaten/kota. Dengan demikian dapat dikatakan bahwa 44.7% kabupaten/kota di Jawa Timur memerlukan perhatian lebih dari segi pembangunan kualitas SDM dan pembangunan ekonomi.

5.2 Saran

Saran yang dapat diberikan berdasarkan analisis dan pembahasan yang telah dilakukan adalah apabila ingin melakukan pengelompokan data campuran dengan topik ini sebaiknya menggunakan data persentase jika tersedia. Hal ini dikarenakan, data persentase dianggap lebih representatif jika dibandingkan dengan data jumlahan. Diharapkan dengan penggunaan data tersebut bisa menghasilkan hasil pengelompokan yang lebih baik.

Selain itu, sebaiknya pemerintah lebih mengoptimalkan adanya kebijakan otonomi daerah sehingga setiap pemimpin daerah dapat mengembangkan potensi ekonomi yang ada di daerahnya guna meningkatkan kesejahteraan masyarakat. Berkembangnya potensi ekonomi dapat meningkatkan sektor investasi di daerahnya

sehingga standar hidup layak masyarakat juga semakin membaik. Selain itu, pembangunan kualitas SDM mutlak perlu dilakukan untuk mendukung berkembangnya potensi ekonomi daerah. Dengan demikian, adanya bonus demografi juga dapat dijadikan sebagai *the windows opportunity*.

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Alvionita. (2017). *Metode Ensembel ROCK Dan SWFM Untuk Pengelompokan Data Campuran Numerik Dan Kategorik Pada Kasus Akses Jeruk*. Tesis. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2012). *Statistics for Business and Economics, Revised Eleventh Edition*. South Western: Cengage Learning.
- Arianti, P. S. (2009). *Pengelompokan Kecamatan di Kabupaten Probolinggo Berdasarkan Indeks Pembangunan Manusia dengan Cluster Analysis*. Tugas Akhir. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- BAPPEDA. (2017). *Data Dinamis Provinsi Jawa Timur Triwulan I - 2017*. Surabaya: Badan Perencanaan Pembangunan Daerah Provinsi Jawa Timur.
- Bunkers, M. J., & James, R. M. (1996). *Definition of Climate Regions in The Northern Plains Using an Objective Cluster Modification Technique*.
- Dewi, A. (2012). *Metode Cluster Ensemble Untuk Pengelompokan Desa Perdesaan di Provinsi Riau*. Tesis. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Dutta, M., Mahanta, A. K., & Arun, K. P. (2005). QROCK : A Quick Version of the ROCK Algorithm for Clustering of Categorical Data. *Proceedings of the 15 IEEE Internasional Conference on Data Engineering*.
- Guha, S., Rastogi, R., & Shim, K. (2000). A Robust Clustering Algorithm for Categorical Attributes. *Proceedings of the 15th IEEE Internasional Conference on Data Engineering*.
- Harmanto, M. N. (2016). *Pengelompokan Mengenai Perubahan Struktur Kependudukan Dalam Menghadapi Era Bonus Demografi di Kabupaten/Kota Provinsi Jawa Timur*. Tugas Akhir. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.

- He, Z., Xu, X., & Deng, S. (2005). *Clustering Mixed Numeric and Categorical Data : A Cluster Ensemble Approach*. Harbin Institute of Technology: Departemen of Computer Science and Engineering.
- Ihsanuddin. (2017, 12 27). *Jokowi : Mulai 2019, Kita Konsentrasi Pada Pembangunan SDM*. Diambil kembali dari KOMPAS.com:
<http://nasional.kompas.com/read/2017/12/27/13195931/jokowi-mulai-2019-kita-konsentrasi-pada-pembangunan-sdm>
- Irfandi, Y. P. (2016). *Penerapan Algoritma Squeezer untuk Pengelompokan Tanaman Jeruk Hasil Fusi Berdasarkan Karakteristik Kualitatif dan Kuantitatif*. Tesis. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). United States of America: Pearson Education Inc.
- Kader, D. G., & Perry, M. (2007). Variability for Categorical Variables. *Journal of Statistics Education*, Vol.15, No.2.
- Lailiya, A. R. (2011). *Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Kesamaan Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka dengan Metode Hierarki dan Non Hierarki*. Tugas Akhir. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Light, R. J., & Margolin, B. H. (1971). An Analysis of Variance for Categorical Data. *Journal of American Statistical Association*, Vol.66 No.335.
- Okada, T. (1999). Sum of Square Decomposition for Categorical Data. *Kwansei Gakuin Studies in Computer Science*, Vol.14.
- Permata, P. I. (2017). *Pengelompokan Perguruan Tinggi Negeri di Indonesia Menggunakan Metode Ensemble ROBust Clustering using linKs (ROCK)*. Tugas Akhir. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.

- Pravitasari, A. A. (2008). *Analisis Pengelompokan dengan Fuzzy C-Means Cluster (Kasus Pengelompokan Kecamatan di Kabupaten Tuban Berdasarkan Tingkat Partisipasi Pendidikan)*. Tesis. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Putra, D. M. (2015). *Pemodelan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur Dengan Menggunakan Metode Regresi Logistik Ridge. Tugas Akhir*. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Satria. (2017, Agustus 30). *Mengoptimalkan Bonus Demografi*. Diambil kembali dari Universitas Gajah Mada: <https://ugm.ac.id/id/berita/14623-mengoptimalkan.bonus.demografi>
- Sharma, S. (1996). *Applied Multivariate Technique* . New York: John Wiley & Sons, Inc. .
- Suguna, J., & Selvi, M. A. (2012). Ensemble Fuzzy Clustering for Mixed Numerical and Categorical Data. *International Journal of Computer Application*, 42, 19.
- Sukirno, S. (2012). *Pengantar Teori Makroekonomi. Edisi 3 Cetakan ke-21*. Jakarta: PT Raja Grafindo Persada.
- Tyagi, A., & Sharma, S. (2012). Implementation of ROCK Clustering Algorithm for the Optimazation of Query Searching Time. *International Journal on Computer Science and Engineering*, Vol. 4 No. 05.
- UNDP. (2017, 12 13). *Sustainable Development Goals* . Diambil kembali dari United Nations Development Programme : <http://www.undp.org/content/undp/en/home/sustainable-development-goals.html>
- UNFPA. (2010). *Taking Advantage of The Demographic Bonus in Viet Nam* . Viet Nam: Ha Noi, United Nations Viet Nam.
- Yoon, H. S., Ahn, S. Y., Lee, S. H., Cho, S. B., & Kim, J. H. (2006). Heterogenous Clustering Ensemble Method for Combining Different Cluster Results. *BioDM 2006*, 82-91.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1 Data Berskala Numerik

No.	Kabupaten/Kota	X1	X2	X3	X4	...	X14	X15	X16	X17
1	Kabupaten Pacitan	3	24	86	61.9	...	87.76	81.82	6.89	1283000
2	Kabupaten Ponorogo	6	31	72	57.4	...	94.65	91.79	6.97	1283000
3	Kabupaten Trenggalek	4	22	82	33.3	...	77.25	84.35	7.19	1283000
4	Kabupaten Tulungagung	12	31	95	42.5	...	93.3	87.7	7.73	1420000
5	Kabupaten Blitar	10	24	88	54	...	95.21	81.06	7.25	1405000
6	Kabupaten Kediri	11	37	84	51.1	...	88.46	80.96	7.58	1456000
7	Kabupaten Malang	24	39	94	30	...	98.96	67.76	6.98	2188000
8	Kabupaten Lumajang	6	25	94	34.8	...	87.58	58.78	6.05	1437000
9	Kabupaten Jember	12	49	91	68	...	93.65	65.42	6.05	1629000
10	Kabupaten Banyuwangi	13	45	80	45.5	...	97.81	79.13	6.93	1599000
11	Kabupaten Bondowoso	4	25	95	19.4	...	100.5	67.01	5.54	1417000
12	Kabupaten Situbondo	5	17	86	25	...	77.99	81.61	5.68	1374000
13	Kabupaten Probolinggo	6	33	94	22.2	...	70.96	43.07	5.67	1736000

No.	Kabupaten/Kota	X1	X2	X3	X4	...	X14	X15	X16	X17
14	Kabupaten Pasuruan	6	33	92	44.5	...	89.56	60.22	6.58	3037500
15	Kabupaten Sidoarjo	26	26	65	62.9	...	101.57	102.07	10.22	3040000
16	Kabupaten Mojokerto	11	27	80	61.9	...	93.43	99.64	7.76	3030000
17	Kabupaten Jombang	13	34	88	46.4	...	87.61	85.46	7.68	1924000
18	Kabupaten Nganjuk	6	20	94	45.8	...	91.07	83.45	7.34	1411000
19	Kabupaten Madiun	4	26	95	71	...	87.43	101.17	7	1340000
20	Kabupaten Magetan	6	22	85	63.9	...	97.99	101.6	7.66	1283000
21	Kabupaten Ngawi	3	24	82	42.5	...	95.85	84.74	6.54	1334000
22	Kabupaten Bojonegoro	10	36	87	59.3	...	90.19	88.79	6.65	1462000
23	Kabupaten Tuban	4	33	81	44.1	...	98.59	78.79	6.25	1757000
24	Kabupaten Lamongan	11	33	89	64	...	90.69	97.61	7.29	1573000
25	Kabupaten Gresik	14	32	58	63.2	...	95.53	112.25	8.94	3042500
26	Kabupaten Bangkalan	4	22	85	71.2	...	87.88	47.88	5.13	1414000
27	Kabupaten Sampang	1	21	95	34.9	...	91.6	54.05	3.79	1387000
28	Kabupaten Pamekasan	5	20	75	34.4	...	90.35	86.87	6.08	1350000
29	Kabupaten Sumenep	3	30	100	55	...	90.15	94.57	5.08	1398000

No.	Kabupaten/Kota	X1	X2	X3	X4	...	X14	X15	X16	X17
31	Kota Blitar	5	3	93	42.5	...	87.57	124.22	9.88	1394000
32	Kota Malang	24	15	82	49.5	...	89.45	77.69	10.14	2099000
33	Kota Probolinggo	4	6	90	63	...	95.83	114.54	8.47	1603000
34	Kota Pasuruan	1	8	79	40.3	...	74.63	111.67	9.08	1757000
35	Kota Mojokerto	7	5	87	59.9	...	90.95	139.87	9.93	1603000
36	Kota Madiun	8	6	83	63.9	...	88.22	113.96	11.09	1394000
37	Kota Surabaya	62	62	76	75.1	...	101.38	90.71	10.44	3045000
38	Kota Batu	5	5	87	28.7	...	94.27	131.06	8.45	2026000

Lampiran 2 Data Berskala Kategori

No.	Kabupaten/Kota	X18	X19	X20	X21
1	Kabupaten Pacitan	1	0	3	3
2	Kabupaten Ponorogo	1	0	3	3
3	Kabupaten Trenggalek	0	0	3	3
4	Kabupaten Tulungagung	1	0	2	3
5	Kabupaten Blitar	1	0	2	3
6	Kabupaten Kediri	1	0	2	3
7	Kabupaten Malang	1	1	1	1
8	Kabupaten Lumajang	1	0	3	3
9	Kabupaten Jember	1	0	2	3
10	Kabupaten Banyuwangi	1	1	2	3
11	Kabupaten Bondowoso	1	0	4	3
12	Kabupaten Situbondo	1	0	4	3
13	Kabupaten Probolinggo	1	0	3	3
14	Kabupaten Pasuruan	1	1	3	1
15	Kabupaten Sidoarjo	1	1	1	1
16	Kabupaten Mojokerto	1	1	2	1
17	Kabupaten Jombang	1	1	2	2
18	Kabupaten Nganjuk	1	0	3	3
19	Kabupaten Madiun	1	0	3	3
20	Kabupaten Magetan	1	1	2	3
21	Kabupaten Ngawi	1	0	3	3
22	Kabupaten Bojonegoro	1	1	3	4
23	Kabupaten Tuban	1	0	3	1
24	Kabupaten Lamongan	1	1	3	3
25	Kabupaten Gresik	1	1	2	1
26	Kabupaten Bangkalan	1	0	4	3

No.	Kabupaten/Kota	X18	X19	X20	X21
27	Kabupaten Sampang	1	1	4	3
28	Kabupaten Pamekasan	1	1	4	3
29	Kabupaten Sumenep	1	0	3	3
30	Kota Kediri	1	1	1	1
31	Kota Blitar	0	1	2	2
32	Kota Malang	0	1	1	2
33	Kota Probolinggo	1	1	3	2
34	Kota Pasuruan	0	1	3	2
35	Kota Mojokerto	1	1	2	2
36	Kota Madiun	0	1	2	2
37	Kota Surabaya	1	1	1	2
38	Kota Batu	1	1	2	2

Lampiran 3 Syntax untuk Menghitung Nilai *Icdrate*

```

Data1 = read.csv("H://Single-4.csv", header=TRUE, sep=";")
Data2 = read.csv("H://Single-3.csv", header=TRUE, sep=";")
Data3 = read.csv("H://Single-2.csv", header=TRUE, sep=";")
Data4 = read.csv("H://Average-4.csv", header=TRUE, sep=";")
Data5 = read.csv("H://Average-3.csv", header=TRUE, sep=";")
Data6 = read.csv("H://Average-2.csv", header=TRUE, sep=";")
Data7 = read.csv("H://Complete-4.csv", header=TRUE, sep=";")
Data8 = read.csv("H://Complete-3.csv", header=TRUE, sep=";")
Data9 = read.csv("H://Complete-2.csv", header=TRUE, sep=";")
Data10 = read.csv("H://Ward-4.csv", header=TRUE, sep=";")
Data11 = read.csv("H://Ward-3.csv", header=TRUE, sep=";")
Data12 = read.csv("H://Ward-2.csv", header=TRUE, sep=";")

icdrate = function(Data, nc)
{
  n = dim(Data)[1]
  p = dim(Data)[2]
  X = Data[,1:(p-1)]
  Group = Data[,p]

  p = dim(X)[2]
  Mean.X = matrix(ncol = p, nrow = (nc+1))
  for (i in 1:nc)
  {
    for (j in 1:p)
    {
      Mean.X[i,j] = mean(X[which(Group==i),j])
      Mean.X[(nc+1),j] = mean(X[,j])
    }
  }

  SST = matrix(ncol=p, nrow=n)
  for (i in 1:n)
  {

```

```

    for (j in 1:p)
    {
        SST[i,j] = (X[i,j] - Mean.X[(nc+1),j])^2
    }
}
SST = sum(sum(SST))
SSE = matrix(ncol=p, nrow=n)
for (i in 1:n)
{
    for (j in 1:p)
    {
        for (k in 1:nc)
        {
            if (Group[i]==k)
            {
                SSE[i,j] = (X[i,j] - Mean.X[k,j])^2
            }
        }
    }
}
SSE = sum(sum(SSE))
Rsqr = (SST-SSE)/SST
icdrate = 1-Rsqr
list(Rsqr=Rsqr, icdrate=icdrate)
}

icdrate(Data1, 4)
icdrate(Data2, 3)
icdrate(Data3, 2)
icdrate(Data4, 4)
icdrate(Data5, 3)
icdrate(Data6, 2)
icdrate(Data7, 4)
icdrate(Data8, 3)
icdrate(Data9, 2)
icdrate(Data10, 4)
icdrate(Data11, 3)
icdrate(Data12, 2)

```

Lampiran 4 Syntax Pengelompokan Data Kategori

```

Data = read.csv("H://Data.csv", header=TRUE, sep=";")

library(dummies)
library(cba)
library(plyr)
library(stats)

Analisis.ROCK = function(Data){
  Datakategori1 = c(Data$X21, Data$X22, Data$X23, Data$X24)
  Datakategori2 = matrix(Datakategori1,38,4)
  X = as.dummy(Datakategori2)
  rc.05 = rockCluster(X, n=2, theta=0.05, debug=FALSE)
  rc.07 = rockCluster(X, n=2, theta=0.07, debug=FALSE)
  rc.10 = rockCluster(X, n=2, theta=0.1, debug=FALSE)
  rc.12 = rockCluster(X, n=2, theta=0.12, debug=FALSE)
  rc.15 = rockCluster(X, n=2, theta=0.15, debug=FALSE)
  rc.17 = rockCluster(X, n=2, theta=0.17, debug=FALSE)
  rc.20 = rockCluster(X, n=2, theta=0.2, debug=FALSE)
  rc.22 = rockCluster(X, n=2, theta=0.22, debug=FALSE)
  rc.25 = rockCluster(X, n=2, theta=0.25, debug=FALSE)
  rc.27 = rockCluster(X, n=2, theta=0.27, debug=FALSE)
  rc.30 = rockCluster(X, n=2, theta=0.30, debug=FALSE)
  rf.05 = fitted(rc.05)
  rf.07 = fitted(rc.07)
  rf.10 = fitted(rc.10)
  rf.12 = fitted(rc.12)
  rf.15 = fitted(rc.15)
  rf.17 = fitted(rc.17)
  rf.20 = fitted(rc.20)
  rf.22 = fitted(rc.22)
  rf.25 = fitted(rc.25)
  rf.27 = fitted(rc.27)
  rf.30 = fitted(rc.30)
  theta.05 = rf.05$c1
  theta.07 = rf.07$c1
  theta.10 = rf.10$c1
  theta.12 = rf.12$c1

```

```

theta.05 = rf.05$cl
theta.07 = rf.07$cl
theta.10 = rf.10$cl
theta.12 = rf.12$cl
theta.15 = rf.15$cl
theta.17 = rf.17$cl
theta.20 = rf.20$cl
theta.22 = rf.22$cl
theta.25 = rf.25$cl
theta.27 = rf.27$cl
theta.30 = rf.30$cl

```

```

Kelompok = data.frame(theta.05, theta.07, theta.10, theta.12, theta.15,
theta.17, theta.20, theta.22, theta.25, theta.27, theta.30)
Hasil = data.frame(Data[2], Kelompok)
write.csv(Hasil, "H:/HasilKlaster-2.csv")

```

```

Jumlah.05 = count(Hasil, "theta.05")
Jumlah.07 = count(Hasil, "theta.07")
Jumlah.10 = count(Hasil, "theta.10")
Jumlah.12 = count(Hasil, "theta.12")
Jumlah.15 = count(Hasil, "theta.15")
Jumlah.17 = count(Hasil, "theta.17")
Jumlah.20 = count(Hasil, "theta.20")
Jumlah.22 = count(Hasil, "theta.22")
Jumlah.25 = count(Hasil, "theta.25")
Jumlah.27 = count(Hasil, "theta.27")
Jumlah.30 = count(Hasil, "theta.30")
X21 = factor (Data$X21)
X22 = factor (Data$X22)
X23 = factor (Data$X23)
X24 = factor (Data$X24)
numtheta.05 = as.numeric(as.character(theta.05))
numtheta.07 = as.numeric(as.character(theta.07))
numtheta.10 = as.numeric(as.character(theta.10))
numtheta.12 = as.numeric(as.character(theta.12))
numtheta.15 = as.numeric(as.character(theta.15))
numtheta.17 = as.numeric(as.character(theta.17))

```

```

numtheta.20 = as.numeric(as.character(theta.20))
numtheta.22 = as.numeric(as.character(theta.22))
numtheta.25 = as.numeric(as.character(theta.25))
numtheta.27 = as.numeric(as.character(theta.27))
numtheta.30 = as.numeric(as.character(theta.30))

```

```

p.05 = numtheta.05~X21+X22+X23+X24
p.07 = numtheta.07~X21+X22+X23+X24
p.10 = numtheta.10~X21+X22+X23+X24
p.12 = numtheta.12~X21+X22+X23+X24
p.15 = numtheta.15~X21+X22+X23+X24
p.17 = numtheta.17~X21+X22+X23+X24
p.20 = numtheta.20~X21+X22+X23+X24
p.22 = numtheta.22~X21+X22+X23+X24
p.25 = numtheta.25~X21+X22+X23+X24
p.27 = numtheta.27~X21+X22+X23+X24
p.30 = numtheta.30~X21+X22+X23+X24

```

```

mylogit.05 = aov(p.05)
mylogit.07 = aov(p.07)
mylogit.10 = aov(p.10)
mylogit.12 = aov(p.12)
mylogit.15 = aov(p.15)
mylogit.17 = aov(p.17)
mylogit.20 = aov(p.20)
mylogit.22 = aov(p.22)
mylogit.25 = aov(p.25)
mylogit.27 = aov(p.27)
mylogit.30 = aov(p.30)

```

```

SSW.05 = sum((mylogit.05$residuals)^2)
SSW.07 = sum((mylogit.07$residuals)^2)
SSW.10 = sum((mylogit.10$residuals)^2)
SSW.12 = sum((mylogit.12$residuals)^2)
SSW.15 = sum((mylogit.15$residuals)^2)
SSW.17 = sum((mylogit.17$residuals)^2)
SSW.20 = sum((mylogit.20$residuals)^2)
SSW.22 = sum((mylogit.22$residuals)^2)

```



```
SSW.25 = sum((mylogit.25$residuals)^2)
SSW.27 = sum((mylogit.27$residuals)^2)
SSW.30 = sum((mylogit.30$residuals)^2)
```

```
h.05 = unlist(summary(mylogit.05))
h.07 = unlist(summary(mylogit.07))
h.10 = unlist(summary(mylogit.10))
h.12 = unlist(summary(mylogit.12))
h.15 = unlist(summary(mylogit.15))
h.17 = unlist(summary(mylogit.17))
h.20 = unlist(summary(mylogit.20))
h.22 = unlist(summary(mylogit.22))
h.25 = unlist(summary(mylogit.25))
h.27 = unlist(summary(mylogit.27))
h.30 = unlist(summary(mylogit.30))
```

```
summary.05 = matrix(h.05, 5,5)
summary.07 = matrix(h.07, 5,5)
summary.10 = matrix(h.10, 5,5)
summary.12 = matrix(h.12, 5,5)
summary.15 = matrix(h.15, 5,5)
summary.17 = matrix(h.17, 5,5)
summary.20 = matrix(h.20, 5,5)
summary.22 = matrix(h.22, 5,5)
summary.25 = matrix(h.25, 5,5)
summary.27 = matrix(h.27, 5,5)
summary.30 = matrix(h.30, 5,5)
```

```
SSB.05 = sum(summary.05[1:4,2])
SSB.07 = sum(summary.07[1:4,2])
SSB.10 = sum(summary.10[1:4,2])
SSB.12 = sum(summary.12[1:4,2])
SSB.15 = sum(summary.15[1:4,2])
SSB.17 = sum(summary.17[1:4,2])
SSB.20 = sum(summary.20[1:4,2])
SSB.22 = sum(summary.22[1:4,2])
SSB.25 = sum(summary.25[1:4,2])
SSB.27 = sum(summary.27[1:4,2])
SSB.30 = sum(summary.30[1:4,2])
```

```

a = sum(rf.05$size>0, na.rm=TRUE)
b = sum(rf.07$size>0, na.rm=TRUE)
c = sum(rf.10$size>0, na.rm=TRUE)
d = sum(rf.12$size>0, na.rm=TRUE)
e = sum(rf.15$size>0, na.rm=TRUE)
f = sum(rf.17$size>0, na.rm=TRUE)
g = sum(rf.20$size>0, na.rm=TRUE)
h = sum(rf.22$size>0, na.rm=TRUE)
i = sum(rf.25$size>0, na.rm=TRUE)
j = sum(rf.27$size>0, na.rm=TRUE)
k = sum(rf.30$size>0, na.rm=TRUE)

```

```

SW.05 = sqrt(SSW.05/(38-a))
SW.07 = sqrt(SSW.07/(38-b))
SW.10 = sqrt(SSW.10/(38-c))
SW.12 = sqrt(SSW.12/(38-d))
SW.15 = sqrt(SSW.15/(38-e))
SW.17 = sqrt(SSW.17/(38-f))
SW.20 = sqrt(SSW.20/(38-g))
SW.22 = sqrt(SSW.22/(38-h))
SW.25 = sqrt(SSW.25/(38-i))
SW.27 = sqrt(SSW.27/(38-j))
SW.30 = sqrt(SSW.30/(38-k))

```

```

SB.05 = sqrt(SSB.05/(a-1))
SB.07 = sqrt(SSB.07/(b-1))
SB.10 = sqrt(SSB.10/(c-1))
SB.12 = sqrt(SSB.12/(d-1))
SB.15 = sqrt(SSB.15/(e-1))
SB.17 = sqrt(SSB.17/(f-1))
SB.20 = sqrt(SSB.20/(g-1))
SB.22 = sqrt(SSB.22/(h-1))
SB.25 = sqrt(SSB.25/(i-1))
SB.27 = sqrt(SSB.27/(j-1))
SB.30 = sqrt(SSB.30/(k-1))

```

```

Ratio.05 = SW.05/SB.05
Ratio.07 = SW.07/SB.07

```

```

Ratio.10 = SW.10/SB.10
Ratio.12 = SW.12/SB.12
Ratio.15 = SW.15/SB.15
Ratio.17 = SW.17/SB.17
Ratio.20 = SW.20/SB.20
Ratio.22 = SW.22/SB.22
Ratio.25 = SW.25/SB.25
Ratio.27 = SW.27/SB.27
Ratio.30 = SW.30/SB.30

Ratio.Kategori = rbind(Ratio.05, Ratio.07, Ratio.10, Ratio.12,
Ratio.15, Ratio.17, Ratio.20, Ratio.22, Ratio.25, Ratio.27, Ratio.30)
print("Hasil Pengelompokan Data Kategori")
print(Hasil)
print("=====
=====")
print("Jumlah Kelompok Yang Terbentuk")
print(Jumlah.05)
print(Jumlah.07)
print(Jumlah.10)
print(Jumlah.12)
print(Jumlah.15)
print(Jumlah.17)
print(Jumlah.20)
print(Jumlah.22)
print(Jumlah.25)
print(Jumlah.27)
print(Jumlah.30)
print(Ratio.05)
print(Ratio.07)
print(Ratio.10)
print(Ratio.12)
print(Ratio.15)
print(Ratio.17)
print(Ratio.20)
print(Ratio.22)
print(Ratio.25)
print(Ratio.27)
print(Ratio.30)

```

```
print("Nilai Ratio Yang Terbentuk")
print(Ratio.Kategori)
aa = c(0.05, 0.07, 0.10, 0.12, 0.15, 0.17, 0.20, 0.22, 0.25, 0.27, 0.30)
bb = c(Ratio.05, Ratio.07, Ratio.10, Ratio.12, Ratio.15, Ratio.17,
Ratio.20, Ratio.22, Ratio.25, Ratio.27, Ratio.30)
plot(aa, bb, main="Plot Untuk Nilai Ratio Metode ROCK",
xlab="Nilai Theta", ylab="Nilai Ratio")
lines(aa, bb, col="Red")
}

Analisis.ROCK(Data)
```

Lampiran 5 Syntax Pengelompokan Data Campuran

```
Data = read.csv("H://Asli-Standardize.csv", header=TRUE, sep=";")

library(dummies)
library(cba)
library(plyr)
library(stats)
Analisis.EnsembleROCK = function(Data){
  Datakategori1 = c(Data$A, Data$B)
  Datakategori2 = matrix(Datakategori1,38,2)
  X = as.dummy(Datakategori2)
  rc.05 = rockCluster(X, n=2, theta=0.05, debug=FALSE)
  rc.07 = rockCluster(X, n=2, theta=0.07, debug=FALSE)
  rc.10 = rockCluster(X, n=2, theta=0.1, debug=FALSE)
  rc.12 = rockCluster(X, n=2, theta=0.12, debug=FALSE)
  rc.15 = rockCluster(X, n=2, theta=0.15, debug=FALSE)
  rc.17 = rockCluster(X, n=2, theta=0.17, debug=FALSE)
  rc.20 = rockCluster(X, n=2, theta=0.2, debug=FALSE)
  rc.22 = rockCluster(X, n=2, theta=0.22, debug=FALSE)
  rc.25 = rockCluster(X, n=2, theta=0.25, debug=FALSE)
  rc.27 = rockCluster(X, n=2, theta=0.27, debug=FALSE)
  rc.30 = rockCluster(X, n=2, theta=0.30, debug=FALSE)
  rf.05 = fitted(rc.05)
  rf.07 = fitted(rc.07)
  rf.10 = fitted(rc.10)
  rf.12 = fitted(rc.12)
  rf.15 = fitted(rc.15)
  rf.17 = fitted(rc.17)
  rf.20 = fitted(rc.20)
  rf.22 = fitted(rc.22)
  rf.25 = fitted(rc.25)
  rf.27 = fitted(rc.27)
  rf.30 = fitted(rc.30)
  theta.05 = rf.05$cl
  theta.07 = rf.07$cl
  theta.10 = rf.10$cl
  theta.12 = rf.12$cl
```

```
theta.15 = rf.15$cl
theta.17 = rf.17$cl
theta.20 = rf.20$cl
theta.22 = rf.22$cl
theta.25 = rf.25$cl
theta.27 = rf.27$cl
theta.30 = rf.30$cl
```

```
Kelompok = data.frame(theta.05, theta.07, theta.10, theta.12,
theta.15, theta.17, theta.20, theta.22, theta.25, theta.27, theta.30)
Hasil = data.frame(Data[2], Kelompok)
write.csv(Hasil, "H:/KlasterEnsembleAsli-2.csv")
```

```
Jumlah.05 = count(Hasil, "theta.05")
Jumlah.07 = count(Hasil, "theta.07")
Jumlah.10 = count(Hasil, "theta.10")
Jumlah.12 = count(Hasil, "theta.12")
Jumlah.15 = count(Hasil, "theta.15")
Jumlah.17 = count(Hasil, "theta.17")
Jumlah.20 = count(Hasil, "theta.20")
Jumlah.22 = count(Hasil, "theta.22")
Jumlah.25 = count(Hasil, "theta.25")
Jumlah.27 = count(Hasil, "theta.27")
Jumlah.30 = count(Hasil, "theta.30")
A = factor(Data$A)
B = factor(Data$B)
```

```
numtheta.05 = as.numeric(as.character(theta.05))
numtheta.07 = as.numeric(as.character(theta.07))
numtheta.10 = as.numeric(as.character(theta.10))
numtheta.12 = as.numeric(as.character(theta.12))
numtheta.15 = as.numeric(as.character(theta.15))
numtheta.17 = as.numeric(as.character(theta.17))
numtheta.20 = as.numeric(as.character(theta.20))
numtheta.22 = as.numeric(as.character(theta.22))
numtheta.25 = as.numeric(as.character(theta.25))
numtheta.27 = as.numeric(as.character(theta.27))
numtheta.30 = as.numeric(as.character(theta.30))
```

```

p.05 = numtheta.05~A+B
p.07 = numtheta.07~A+B
p.10 = numtheta.10~A+B
p.12 = numtheta.12~A+B
p.15 = numtheta.15~A+B
p.17 = numtheta.17~A+B
p.20 = numtheta.20~A+B
p.22 = numtheta.22~A+B
p.25 = numtheta.25~A+B
p.27 = numtheta.27~A+B
p.30 = numtheta.30~A+B

```

```

mylogit.05 = aov(p.05)
mylogit.07 = aov(p.07)
mylogit.10 = aov(p.10)
mylogit.12 = aov(p.12)
mylogit.15 = aov(p.15)
mylogit.17 = aov(p.17)
mylogit.20 = aov(p.20)
mylogit.22 = aov(p.22)
mylogit.25 = aov(p.25)
mylogit.27 = aov(p.27)
mylogit.30 = aov(p.30)

```

```

SSW.05 = sum((mylogit.05$residuals)^2)
SSW.07 = sum((mylogit.07$residuals)^2)
SSW.10 = sum((mylogit.10$residuals)^2)
SSW.12 = sum((mylogit.12$residuals)^2)
SSW.15 = sum((mylogit.15$residuals)^2)
SSW.17 = sum((mylogit.17$residuals)^2)
SSW.20 = sum((mylogit.20$residuals)^2)
SSW.22 = sum((mylogit.22$residuals)^2)
SSW.25 = sum((mylogit.25$residuals)^2)
SSW.27 = sum((mylogit.27$residuals)^2)
SSW.30 = sum((mylogit.30$residuals)^2)

```

```

h.05 = unlist(summary(mylogit.05))
h.07 = unlist(summary(mylogit.07))

```

```
h.10 = unlist(summary(mylogit.10))
h.12 = unlist(summary(mylogit.12))
h.15 = unlist(summary(mylogit.15))
h.17 = unlist(summary(mylogit.17))
h.20 = unlist(summary(mylogit.20))
h.22 = unlist(summary(mylogit.22))
h.25 = unlist(summary(mylogit.25))
h.27 = unlist(summary(mylogit.27))
h.30 = unlist(summary(mylogit.30))
```

```
summary.05 = matrix(h.05, 3,3)
summary.07 = matrix(h.07, 3,3)
summary.10 = matrix(h.10, 3,3)
summary.12 = matrix(h.12, 3,3)
summary.15 = matrix(h.15, 3,3)
summary.17 = matrix(h.17, 3,3)
summary.20 = matrix(h.20, 3,3)
summary.22 = matrix(h.22, 3,3)
summary.25 = matrix(h.25, 3,3)
summary.27 = matrix(h.27, 3,3)
summary.30 = matrix(h.30, 3,3)
```

```
SSB.05 = sum(summary.05[1:2,2])
SSB.07 = sum(summary.07[1:2,2])
SSB.10 = sum(summary.10[1:2,2])
SSB.12 = sum(summary.12[1:2,2])
SSB.15 = sum(summary.15[1:2,2])
SSB.17 = sum(summary.17[1:2,2])
SSB.20 = sum(summary.20[1:2,2])
SSB.22 = sum(summary.22[1:2,2])
SSB.25 = sum(summary.25[1:2,2])
SSB.27 = sum(summary.27[1:2,2])
SSB.30 = sum(summary.30[1:2,2])
```

```
a = sum(rf.05$size>0, na.rm=TRUE)
b = sum(rf.07$size>0, na.rm=TRUE)
c = sum(rf.10$size>0, na.rm=TRUE)
d = sum(rf.12$size>0, na.rm=TRUE)
```



```
e = sum(rf.15$size>0, na.rm=TRUE)
f = sum(rf.17$size>0, na.rm=TRUE)
g = sum(rf.20$size>0, na.rm=TRUE)
h = sum(rf.22$size>0, na.rm=TRUE)
i = sum(rf.25$size>0, na.rm=TRUE)
j = sum(rf.27$size>0, na.rm=TRUE)
k = sum(rf.30$size>0, na.rm=TRUE)
```

```
SW.05 = sqrt(SSW.05/(38-a))
SW.07 = sqrt(SSW.07/(38-b))
SW.10 = sqrt(SSW.10/(38-c))
SW.12 = sqrt(SSW.12/(38-d))
SW.15 = sqrt(SSW.15/(38-e))
SW.17 = sqrt(SSW.17/(38-f))
SW.20 = sqrt(SSW.20/(38-g))
SW.22 = sqrt(SSW.22/(38-h))
SW.25 = sqrt(SSW.25/(38-i))
SW.27 = sqrt(SSW.27/(38-j))
SW.30 = sqrt(SSW.30/(38-k))
```

```
SB.05 = sqrt(SSB.05/(a-1))
SB.07 = sqrt(SSB.07/(b-1))
SB.10 = sqrt(SSB.10/(c-1))
SB.12 = sqrt(SSB.12/(d-1))
SB.15 = sqrt(SSB.15/(e-1))
SB.17 = sqrt(SSB.17/(f-1))
SB.20 = sqrt(SSB.20/(g-1))
SB.22 = sqrt(SSB.22/(h-1))
SB.25 = sqrt(SSB.25/(i-1))
SB.27 = sqrt(SSB.27/(j-1))
SB.30 = sqrt(SSB.30/(k-1))
```

```
Ratio.05 = SW.05/SB.05
Ratio.07 = SW.07/SB.07
Ratio.10 = SW.10/SB.10
Ratio.12 = SW.12/SB.12
Ratio.15 = SW.15/SB.15
Ratio.17 = SW.17/SB.17
```

```

Ratio.20 = SW.20/SB.20
Ratio.22 = SW.22/SB.22
Ratio.25 = SW.25/SB.25
Ratio.27 = SW.27/SB.27
Ratio.30 = SW.30/SB.30

Ratio.Kategori = rbind(Ratio.05, Ratio.07, Ratio.10, Ratio.12,
Ratio.15, Ratio.17, Ratio.20, Ratio.22, Ratio.25, Ratio.27, Ratio.30)
print("Hasil Pengelompokan Data Kategori")
print(Hasil)
print("=====
=====")
print("Jumlah Kelompok Yang Terbentuk")
print(Jumlah.05)
print(Jumlah.07)
print(Jumlah.10)
print(Jumlah.12)
print(Jumlah.15)
print(Jumlah.17)
print(Jumlah.20)
print(Jumlah.22)
print(Jumlah.25)
print(Jumlah.27)
print(Jumlah.30)

print(Ratio.05)
print(Ratio.07)
print(Ratio.10)
print(Ratio.12)
print(Ratio.15)
print(Ratio.17)
print(Ratio.20)
print(Ratio.22)
print(Ratio.25)
print(Ratio.27)
print(Ratio.30)
print("Nilai Ratio Yang Terbentuk")
print(Ratio.Kategori)

```

```
aa = c(0.05, 0.07, 0.10, 0.12, 0.15, 0.17, 0.20, 0.22, 0.25, 0.27, 0.30)
bb = c(Ratio.05, Ratio.07, Ratio.10, Ratio.12, Ratio.15, Ratio.17,
Ratio.20, Ratio.22, Ratio.25, Ratio.27, Ratio.30)
plot(aa, bb, main="Plot Untuk Nilai Ratio Metode ROCK",
xlab="Nilai Theta", ylab="Nilai Ratio")
lines(aa, bb, col="Red")
}
```

```
Analisis.EnsembleROCK(Data)
```

Lampiran 6 Anggota Pengelompokan Data Numerik

No.	Kabupaten/Kota	Kelompok Numerik
1	Kabupaten Pacitan	1
2	Kabupaten Ponorogo	1
3	Kabupaten Trenggalek	1
4	Kabupaten Tulungagung	1
5	Kabupaten Blitar	1
6	Kabupaten Kediri	1
7	Kabupaten Malang	1
8	Kabupaten Lumajang	1
9	Kabupaten Jember	2
10	Kabupaten Banyuwangi	1
11	Kabupaten Bondowoso	2
12	Kabupaten Situbondo	2
13	Kabupaten Probolinggo	2
14	Kabupaten Pasuruan	2
15	Kabupaten Sidoarjo	3
16	Kabupaten Mojokerto	1
17	Kabupaten Jombang	1
18	Kabupaten Nganjuk	1
19	Kabupaten Madiun	1
20	Kabupaten Magetan	1
21	Kabupaten Ngawi	1
22	Kabupaten Bojonegoro	1
23	Kabupaten Tuban	1
24	Kabupaten Lamongan	1
25	Kabupaten Gresik	3
26	Kabupaten Bangkalan	2

No.	Kabupaten/Kota	Kelompok Numerik
27	Kabupaten Sampang	2
28	Kabupaten Pamekasan	2
29	Kabupaten Sumenep	2
30	Kota Kediri	4
31	Kota Blitar	4
32	Kota Malang	1
33	Kota Probolinggo	1
34	Kota Pasuruan	4
35	Kota Mojokerto	4
36	Kota Madiun	4
37	Kota Surabaya	3
38	Kota Batu	4

Lampiran 7 Anggota Pengelompokan Data Kategori

No.	Kabupaten/Kota	Kelompok Kategori
1	Kabupaten Pacitan	1
2	Kabupaten Ponorogo	1
3	Kabupaten Trenggalek	1
4	Kabupaten Tulungagung	1
5	Kabupaten Blitar	1
6	Kabupaten Kediri	1
7	Kabupaten Malang	2
8	Kabupaten Lumajang	1
9	Kabupaten Jember	1
10	Kabupaten Banyuwangi	2
11	Kabupaten Bondowoso	1
12	Kabupaten Situbondo	1
13	Kabupaten Probolinggo	1
14	Kabupaten Pasuruan	2
15	Kabupaten Sidoarjo	2
16	Kabupaten Mojokerto	2
17	Kabupaten Jombang	2
18	Kabupaten Nganjuk	1
19	Kabupaten Madiun	1
20	Kabupaten Magetan	2
21	Kabupaten Ngawi	1
22	Kabupaten Bojonegoro	2
23	Kabupaten Tuban	1
24	Kabupaten Lamongan	2
25	Kabupaten Gresik	2
26	Kabupaten Bangkalan	1

No.	Kabupaten/Kota	Kelompok Kategori
27	Kabupaten Sampang	2
28	Kabupaten Pamekasan	2
29	Kabupaten Sumenep	1
30	Kota Kediri	2
31	Kota Blitar	2
32	Kota Malang	2
33	Kota Probolinggo	2
34	Kota Pasuruan	2
35	Kota Mojokerto	2
36	Kota Madiun	2
37	Kota Surabaya	2
38	Kota Batu	2

Lampiran 8 Anggota Pengelompokan Data Campuran

No.	Kabupaten/Kota	Kelompok Campuran
1	Kabupaten Pacitan	1
2	Kabupaten Ponorogo	1
3	Kabupaten Trenggalek	1
4	Kabupaten Tulungagung	1
5	Kabupaten Blitar	1
6	Kabupaten Kediri	1
7	Kabupaten Malang	2
8	Kabupaten Lumajang	1
9	Kabupaten Jember	1
10	Kabupaten Banyuwangi	2
11	Kabupaten Bondowoso	1
12	Kabupaten Situbondo	1
13	Kabupaten Probolinggo	1
14	Kabupaten Pasuruan	2
15	Kabupaten Sidoarjo	2
16	Kabupaten Mojokerto	2
17	Kabupaten Jombang	2
18	Kabupaten Nganjuk	1
19	Kabupaten Madiun	1
20	Kabupaten Magetan	2
21	Kabupaten Ngawi	1
22	Kabupaten Bojonegoro	2
23	Kabupaten Tuban	1
24	Kabupaten Lamongan	2
25	Kabupaten Gresik	2
26	Kabupaten Bangkalan	1

No.	Kabupaten/Kota	Kelompok Campuran
27	Kabupaten Sampang	2
28	Kabupaten Pamekasan	2
29	Kabupaten Sumenep	1
30	Kota Kediri	2
31	Kota Blitar	2
32	Kota Malang	2
33	Kota Probolinggo	2
34	Kota Pasuruan	2
35	Kota Mojokerto	2
36	Kota Madiun	2
37	Kota Surabaya	2
38	Kota Batu	2

Lampiran 9 Hasil Pengelompokan Data Kategori dengan
Threshold Lainnya

<i>Threshold</i>	Jumlah Kelompok	Rasio S_w/S_B
0.01	2	NaN
	3	NaN
	4	NaN
0.02	2	NaN
	3	NaN
	4	NaN
0.03	2	0.236
	3	NaN
	4	NaN
0.4	2	0.045
	3	NaN
	4	NaN
0.5	2	0.045
	3	NaN
	4	NaN
0.6	2	NaN
	3	NaN
	4	NaN
≥ 0.7	2	<i>Error</i>
	3	<i>Error</i>
	4	<i>Error</i>

Keterangan :

NaN menunjukkan bahwa dengan *threshold* tersebut, data tidak dapat dikelompokkan sesuai dengan kelompok yang diminta. Sehingga rasio tidak dapat dihitung.

Lampiran 10 Hasil Pengelompokan Data Campuran dengan
Threshold Lainnya

<i>Threshold</i>	Jumlah Kelompok	Rasio S_w/S_B
0.01	2	0.04832
	3	NaN
	4	NaN
0.02	2	0.04832
	3	NaN
	4	NaN
0.03	2	0.04832
	3	NaN
	4	NaN
0.4	2	NaN
	3	NaN
	4	NaN
0.5	2	NaN
	3	NaN
	4	NaN
0.6	2	NaN
	3	NaN
	4	NaN
≥ 0.7	2	NaN
	3	NaN
	4	NaN

Keterangan :

NaN menunjukkan bahwa dengan *threshold* tersebut, data tidak dapat dikelompokkan sesuai dengan kelompok yang diminta. Sehingga rasio tidak dapat dihitung.

Lampiran 11 Surat Pernyataan Pengambilan Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini adalah mahasiswa Departemen Statistika FMKSD ITS :

Nama : Lilik Indriyati Ichsan

NRP : 0621144000040

Menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari ~~penelitian/ buku/ Tugas Akhir/ Thesis/~~ publikasi lainnya, yaitu :

- Sumber :
1. Peraturan Gubernur Jawa Timur Nomor 68 Tahun 2015
 2. Peraturan Menteri Komunikasi dan Informatika Republik Indonesia Nomor 39 Tahun 2008
 3. Profil Kesehatan Provinsi Jawa Timur Tahun 2016
 4. Website BPS Pusat dan BPS masing-masing kabupaten/kota

- Keterangan :
1. Upah Minimum Kabupaten/Kota (UMK) di Jawa Timur Tahun 2016
 2. Pembagian daerah ekonomi untuk kabupaten/kota
 3. Persentase ketersediaan obat dan vaksin, persentase Rumah Tangga (RT) berperilaku hidup bersih dan sehat
 4. Indikator kesehatan, pendidikan, ekonomi lainnya.

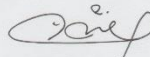
Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data, maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir

Surabaya, 26 Juli 2018



(Erma Oktania Permatasari, S.Si., M.Si)
NIP. 19881007 201404 2 002



(Lilik Indriyati Ichsan)
NRP. 0621144000040

*(Coret yang tidak perlu)

BIODATA PENULIS



Penulis dengan nama lengkap Lilik Indriyati Ichsan merupakan anak pertama dari dua bersaudara. Dilahirkan pada tanggal 04 Mei 1996, di Kabupaten Karanganyar, Provinsi Jawa Tengah. Penulis menempuh pendidikan formal di SDN Pacul 1 Bojonegoro, SMPN 1 Bojonegoro, dan SMA Negeri 1 Bojonegoro. Kemudian penulis melanjutkan pendidikan ke Institut Teknologi Sepuluh Nopember sebagai mahasiswa di Departemen Statistika melalui jalur SNMPTN pada tahun 2014. Selama masa perkuliahan penulis aktif di beberapa kegiatan kepanitiaan, salah satunya adalah sebagai bendahara *Statistics Competition* (STATION) 2016. Selain itu, penulis juga aktif di beberapa organisasi, antara lain sebagai bendahara II HIMASTA-ITS 2015/2016, bendahara I HIMASTA-ITS 2016/2017, staff bisnis KOPMA Dr. Angka ITS 2015/2016, dan asisten bidang bisnis 2016/2017. Selama menjadi mahasiswa penulis diberi kesempatan untuk menjadi semifinalis lomba *National Statistics Competition* (NSC) pada tahun 2017. Kemudian di tahun yang sama penulis juga berkesempatan untuk melakukan kerja praktik di Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN) bidang Keluarga Sejahtera & Pemberdayaan Keluarga. Selain melalui kerja praktik, untuk mengaplikasikan ilmu statistika yang telah didapat selama perkuliahan, penulis juga mengikuti beberapa kegiatan survey. Berkaitan dengan Tugas Akhir ini, apabila pembaca ingin memberikan kritik dan saran, serta diskusi lebih lanjut, pembaca dapat menghubungi penulis melalui *email* berikut lilikindriyati@gmail.com.